# DO DIFFERENT VOICE QUALITIES HAVE THE ABILITY TO EVOKE AN EMOTIONAL RESPONSE WITHIN THE LISTENER?

Michaela Penny

# DO DIFFERENT VOICE QUALITIES HAVE THE ABILITY TO EVOKE AN EMOTIONAL RESPONSE WITHIN THE LISTENER?

By

MICHAELA PENNY

A thesis submitted to the University of Bedfordshire, in fulfillment of the requirements for the Master's degree of Research.

July 2015

**DECLARATION**

I declare that this thesis is my own unaided work. It is being submitted for the degree of Masters of Research at the University of Bedfordshire.

It has not been submitted before for any degree or examination in any other University.

Name of candidate: Michaela Penny Signature:

Date: 6th July 2015

## Contents

1. Introduction ..... 1
2. Plan of Thesis ..... 5
**2.1 The work of Jo Estill** ..... 5
**2.2 Processing emotional responses** ..... 7
**2.3 Factors that influence our emotional responses** ..... 12
**2.4 The difficulties with perception** ..... 17
3. Methodology design ..... 1
**3.1 Implementing the methodology** ..... 17
4. Analysis of the study ..... 30
**4.1 Research Findings** ..... 33
5. Evaluation ..... 1
**5.1 Conclusion** ..... 5
**5.2 Further research** ..... 10
6. Appendices ..... 12
**6.1 Appendix A : Physiology and anatomy of the voice** ..... 13
**6.2 Appendix B : Instructions for the participants** ..... 16
**6.3 Appendix C: Squiggle booklets** ..... 17
**6.4 Appendix D: Description of voice qualities table** ..... 18
**6.5 Appendix E: Colour key chart** ..... 22
7. Appendices reference list ..... 29
8. Thesis reference list ..... 30

# Table of figures and tables

Figure 1- The structures that make up the larynx..... 5
Figure 2- Example of participant booklet..... 23

Table 1- Descriptions of voice qualities……..40
Table 2- Colour key chart……..60
Table 3- Table of opportunites……..61

## Abstract

The purpose of this research is to establish whether different voice qualities evoke an emotional response within the listener. To assist the research, a study was conducted under controlled conditions using eight participants. They were asked to take part in a modified version of Donald Winnicott's squiggle game. This involved allowing the hand to draw freely whilst listening to an original composition for solo voice. The voice qualities demonstrated in the composition were those defined by researcher, voice scientist and teacher, Jo Estill. Participants were informed that a squiggle in an upward direction would indicate an increase in intensity of emotional response. The results revealed that the voice qualities featured in the composition evoked some form of emotional response within the listener. However, an inability to control all variables meant that it was not possible to decipher whether these responses were elicited as a result of voice quality alone.

# 1. Introduction

The focus of this thesis is to answer the question: do different voice qualities have the ability to evoke an emotional response within the listener? The inspiration for this research derives from my experiences as a performer, teacher and listener. I am sure that I speak on behalf of most artists when I state that as a performer, the objective is to move or affect the audience in some way. Although I have seemingly achieved this successfully from time to time, I have never been sure how or why. Initially I assumed that it was the construction of the melody and the message of the lyrical content that made this possible. However, as an audience member who has been emotionally affected by listening to a song, often I am unable to pinpoint what I am feeling and why I am feeling it. Drawing on a specific example, when I first heard Taylor Swift's *I knew you were trouble* (Swift, 2012), I was not particularly moved by the song or the vocal performance, nor did I pay much attention to the lyrical content. In contrast, when I listened to Jessie J's cover of this track (*Radio One Live Lounge*, 2013), I experienced an intensely emotional response to the performance and immediately connected to the message of the song. Although there were adaptations made to the melody in Jessie J's version, the words and the story were exactly the same. This suggests that it was the *way* in which the voice was used in the cover that had an impact on me emotionally. Jessie J executed a fantastically expressive vocal performance that demonstrated the wide range of possibilities the human

voice has to offer. This led me to consider *what* she was doing with her voice and *why* it elicited such a strong reaction within me. In order to explore these questions further, I needed a way of defining these different sounds.

This led me to the work of singer, researcher, voice scientist and teacher, Jo Estill. Detailed information on the work of Jo Estill can be found in *Estill Voice Training Level One: Figures for Voice Control* (Klimek et al., 2005a). Estill began singing at a young age and despite enjoying a successful career as a singer, she often questioned “How am I doing this” (Klimek, et al., 2005a, p.1). Her lack of technical understanding of what she was doing from performance to performance resulted in anxiety over the unpredictability of the outcome[1]. In an attempt to answer her queries surrounding voice production, Estill decided to study a Master’s Degree in Music Education. Through the study of vocal anatomy and physiology, respiration and acoustics, she began to unearth some answers to her persisting question. Years of extensive research into the production of different voice qualities followed, involving x-ray observations, acoustic analysis and perceptual studies. This resulted in the development of the Estill Voice Training Model. Here, the different structures that make up the larynx and vocal tract are identified along with exercises to gain independent control over them. This was all possible due to observations made by Estill regarding the conditions of these different structures during

[1] P.Penny, 2015

the production of different voice qualities. This is to say that the Estill Model, which transformed the world of voice training, began with the study of voice quality. Jo Estill states, “Many who deal with the voice are very much concerned with the ‘quality of voice’. Quality refers to a perceptual attribute of voice and concerns the psychological domain. It is a listener’s perception that the singer, actor, or salesman wishes to influence or affect. It would seem logical, therefore, to begin any study of voice quality by exploring the perceptual domain” (Colton and Estill, 1981, p.323). Colton and Estill’s insights support my hypothesis that voice quality has the ability to evoke a response within the listener and highlights the necessity to explore the role of perception. It became apparent to me that what I was experiencing when listening to Jessie J’s cover was an emotional response to different voice qualities. With information on the production of these different voice qualities outlined in the Estill Model, I could then set about connecting them to the psychological realm.

Following this introduction is the plan of the thesis which also contains a literature review of the following: Jo Estill and her Model; existing research surrounding the connection between musical stimuli and emotion; factors that influence the processing of emotion and descriptions of previous perceptive based studies that highlight different methodologies and potential issues and variables that commonly arise.

Leading on from the plan of thesis is an outline of the design and implementation of the study constructed to address the research question. This includes a breakdown of the different voice qualities (as defined by Jo Estill) that were featured in the vocal composition and formed the focus for the study. Additionally, the research and work of pediatrician and psychoanalyst Donald Winnicott is also described, in particular his squiggle game, which provides the structure for the chosen methodology.

Following the design and implementation of the methodology section are the details of the results and analysis of the study which aimed to discover whether different voice qualities have the ability to evoke an emotional response within the listener. Identification of the different voice qualities featured at different points in the song, alongside analysis of corresponding peaks and troughs of emotional responses are discussed and analysed.

To end the thesis is an evaluation of the study and a conclusion with regards to the research question and ideas for further research.

# 2. Plan of Thesis

This part of the thesis focuses on two separate elements. The first is the innovative research of voice scientist Jo Estill, which will be discussed alongside the insights derived from the relevant literature that is based on the processing of emotional response as a result of musical stimuli. Secondly, the role that perception has on these responses will be used to orientate the research. In addition, previous studies devised by the likes of Krieman (1993) and Jones (2005) are addressed to assist in the development of this study, in order to answer the research question posed at the beginning of this thesis.

## *2.1 The work of Jo Estill*

Detailed information on the work of Jo Estill can be found in *Estill Voice Training Level One: Figures for Voice Control* (Klimek et al., 2005a). In her quest to discover what she was doing with her own voice, Jo Estill embarked on a journey of extensive voice research. After establishing four voice qualities: modal quality (speech), sob, twang and opera, Estill recorded singers producing each quality on different frequencies across the vocal range. Analysis of the acoustic characteristics of each quality followed. Physiology came to play a key role in her understanding of voice quality, with the size and shape of the vocal tract for each of these modes being measured and analysed under x-ray observation. Additionally, endoscopic studies of several subjects followed as well as a perceptual

study, which gave listeners the task of identifying each of the qualities. This topic will be expanded in the methodology section of this thesis, as it is one of the underpinnings of this research.

Further information surrounding Jo Estill's research can be found in *Estill Voice Training System Level two: Figure Combinations for Six Voice Qualities* (Klimek, et al., 2005b). Estill found that, "Seven structures were noted to change position or size during the voice qualities studied" (Klimek, et al., 2005b, p.3). This led her to hypothesis (2005) that each of the seven structures identified could be isolated. Estill later discovered two additional voice qualities, falsetto and belt, which resulted in a total of six qualities that could be utilised in further research. By the beginning of the year 2000, Estill had established thirteen structures along with exercises that could be deployed to gain independent control over them. These are collated in Estill Voice Training Level One: *Figures for Voice Control* and provide the basis for the level one course in Estill Voice Training. *Estill Voice Training System Level Two: Figure Combinations for Six Voice Qualities* developed later, providing the foundation for the level two course in Estill Voice Training. Here, the six voice qualities are presented as combinations of the different structural conditions outlined in *Estill Voice Training Level One: Figures for Voice Control*. The production of some of the voice qualities defined by Jo Estill will be outlined in more detail in the methodology section of this thesis.

It is essential to review what has already been uncovered about the connection between musical stimuli and emotion as part of establishing whether different voice qualities have the ability to evoke an emotional response within the listener. Significantly, at present there is a paucity of published research regarding the emotional response of individuals when listening to different voice qualities in the context of singing. However, select literature that focuses on music and emotion is, nonetheless, worth exploring.

### *2.2 Processing emotional responses*

The following section of the thesis refers to affect and emotion in terms of a response to musical stimulus. The distinction is important: One may feel affected by the stimulus without being able to articulate what emotions have been felt. This is to say that the affect of something can be somewhat overwhelming, whereas the emotion, if felt, might be limited. Juslin and Sloboda (2001) refer to emotions as ‘musical emotions’, suggesting that these may differ from other emotions. While this is a valid point the research in hand is concerned with the potential affect different voice qualities might have on the listener as opposed to the variety of emotions (musical emotions) that might be experienced.

Juslin and Sloboda state “A central aim of psychology is to understand the mechanisms that intervene between music reaching a person’s ears and an emotion being experienced or detected by that person as a result of

hearing music" (2001, p.72). Juslin (2010) and Västfjäll (2008) developed a framework of seven psychological mechanisms known as the BRECVEM model. These are defined as brain stem reflexes, rhythmic entrainment, evaluative conditioning, emotional contagion, visual imagery, episodic memory and musical expectancy (Juslin et al., 2010). These may provide reasoning behind the experience of feeling unexplainably moved by music, or perhaps more specifically, being moved to tears by a vocalist. Below, each mechanism is defined with an interpretation of how it could influence emotional response to different voice qualities:

a) Brain stem reflex could bring about an emotional response due to specific features of the stimulus proceeding to the brain stem to prompt significant information. This refers to acoustic features of the stimulus, such as tempo (Juslin et al., 2010). This could indicate that a high, loud note, (high frequency and high amplitude) produced by a vocalist, could grab the listener's attention prompting feelings such as surprise. This could appear as an urgent event;

b) Rhythmic entrainment could evoke an emotional response due to the rhythmic structure of what is being heard communicating with our body's natural rhythms. (Juslin et al., 2010). This suggests that we might feel

affected when listening to a vocalist due to the rhythmic characteristics utilised (possibly the addition of vibrato) correlating with our biological functions. For example, our heart rate/pulse could adapt to the same rhythm as the stimulus, consequently eliciting a strong response within the listener (Juslin et al., 2010);

c) Evaluative conditioning could elicit an emotional response due to an association made between the stimulus and personal experiences (Juslin et al., 2010). For example, a particular voice quality could serve as a reminder of a particular time, event or person. For somebody who has experienced grief, sob quality could have the power to take them back to that particular time of mourning. Refer to the methodology section of the thesis for information on sob quality;

d) Emotional contagion could exhibit an emotional response due to the listener imitating the emotional state communicated by the performer (Juslin, 2001). For example, if the singer is using cry quality, in mimicking the feelings conveyed by the singer, the listener might also sub-consciously mimic the laryngeal posture of the singer. This is the same posture that the larynx adopts when we

cry for real and could therefore bring about feelings of sadness and possibly even tears. Refer to the methodology section of the thesis for information on cry quality;

e) Visual imagery could result in an emotional response due to imagery created by the listener when exposed to the stimulus (Juslin et al., 2010). For example, if a singer maintains a low larynx position, the lower harmonics are amplified resulting in a dark sound. This perceived acoustic correlate could indeed trigger a specific image in the imagination of the listener. Refer to the methodology section of the thesis for information on lowering the larynx;

f) Episodic memory might awaken an emotional response within the listener by triggering autobiographical episodes from an individual's memory (Baumgartner, 1992). A specific voice quality may arouse a particular memory from the past, consequently evoking the emotional response connected to that memory. "When the memory is evoked, so also is the emotion associated with the memory" (Juslin et al., 2010, p.623);

g) Musical expectancy could bring about an emotional response within the listener due to the stimulus differing from what is expected (Juslin et al., 2010). For example, an

> imperfect cadence might arouse surprise or even confusion in the listener as the unresolved musical phrase opposes common expectation. Perhaps then the unexpected incorporation of distortion in the middle of a refined and aesthetically pleasing vocal performance could evoke a similar emotional response.

These mechanisms present one possible reason behind the emotional responses listeners experience as a result of musical stimuli. It is difficult to distinguish whether one or all of these mechanisms are solely responsible at any given time, or whether it is a combination of several. Furthermore, the question as to whether these individual mechanisms are innate and exist prior to birth or whether they are part of human development proves to be a thought worth considering, even if it is not possible to arrive at a definitive conclusion. Such issues are examined in the work of Juslin et al, who note, “Brain stem reflexes to music could be functional even prior to birth, whereas responses involving musical expectancy may not develop fully until between the ages of 5 to 11” (2010, p.624). This suggests that our musical expectations are influenced by developmental elements such as environmental factors, personal experiences and up-bringing. Therefore, these expectations and consequently the emotional responses associated with them will inevitably differ from individual to individual. In contrast, reflexes to music that are

present in the womb are likely to be consistently inherent from person to person (Juslin et al., 2010).

How conscious we are of our responses to music will unavoidably affect how we articulate them. The processing of emotion is very complex. Often we might feel something yet be unable to describe or pinpoint exactly what that specific feeling is. This could pose some challenges when attempting to measure responses to different voice qualities, as there will be a certain amount of reliance on the ability of participants to communicate their responses clearly.

### *2.3 Factors that influence our emotional responses*

There are many environmental and social factors that influence our emotional response to stimuli. The research that follows focuses on these elements and explores some of the issues surrounding perception based studies. As human beings we experience emotions every single day. Emotion is a huge part of our existence and both effects and affects almost everything we think and do. So influential are our emotions that they are the driving force for artists when devising their work (Juslin and Sloboda, 2001). The artist will often want to convey a particular feeling, or evoke some kind of emotional response within their audience; how effective this is often determines the perceived success of the performance. Even though we might be consciously aware that we experience emotions, nevertheless, our knowledge of them might be

somewhat basic. We know that we experience a range of feelings and that they consist of both good and bad emotions (Juslin and Sloboda, 2001). However, in reality they are much more complicated than this and their complexity is what makes them tricky to articulate and describe.

Juslin and Sloboda (2001) define emotions in relation to music specifically as ‘musical emotions.’ Significantly, these emotions are affected by the simple beauty of the stimulus itself. Secondly, they are influenced by the mood that is implied by the stimulus (Juslin & Sloboda, 2001). These ‘musical emotions’ could well be relevant in the context of singing and different voice qualities. That is to say that the beauty of a singer’s voice provides aesthetic value while the chosen voice quality determines the mood. The mood evoked within the listener has the ability to affect their emotional response and this highlights a separate issue surrounding the influence of taste on emotion. If emotions are initially affected by aesthetic value then this suggests that if listeners do not perceive the voice to sound beautiful or appealing, their initial response might well be one of simple distaste. This was an important element that had to be taken into consideration when structuring this study. In addition, measuring the listener’s response to the mood evoked by different voice qualities is challenging as emotions are perceived differently from individual to individual, partly because our emotions and also our ability to interpret the emotions of others develop over time, as we do (Juslin and Sloboda, 2001).

Listening to different voice qualities might elicit specific emotions, however it is difficult to establish whether these emotions are impulsive, or whether they are highly influenced by our previous experiences or perhaps particular episodic memories. The role that past experiences and associations play in our interpretation of music is acknowledged in the world of music therapy and supported in the work of Bunt and Pavlicevic who comment, "…music is a very powerful trigger that can set off a whole range of associations with specific events and places, memories of certain people, particularly if the experiences or people have been significant in a patient's life" (2001, p.184). The influence of previous experiences on our emotional responses to musical stimuli is not necessarily something that we are consciously aware of. It is quite possible that an individual will express an intense emotional response to music such as crying, while at the same time being unable to articulate or even understand why they feel such a way (Juslin and Sloboda, 2001). This suggests that emotions aroused when listening to music or different voice qualities can be a result of a subconscious reaction with the potential to present different sets of emotions in different people.

It is part of human nature to affiliate past experiences or events with music. Particular sounds or features in a piece of music can evoke specific feelings connected to these events (Bunt and Pavlicevic, 2001). This can include lyrical content which highlights a potential variable for this study. In attempting to assess the influence that different voice qualities have on

emotional reactions, it would be wise to acknowledge the potential impact that the lyrics could have on the results.

Our emotional maturity develops over time, changing the way that we perceive things (Juslin and Sloboda, 2001). Therefore, responses to music and different voice qualities can differ depending on age and life experience. Additionally, it is difficult to decipher if emotions are evoked solely by the built-in characteristics of a piece of music. This is supported in the work of Gregory (1996), “There is no general agreement as to whether these relationships between musical form and emotional feelings are due to the inherent qualities of the music or the learning of associations during early musical experiences” (1996, p.47).

Cultural influences might also affect the listener’s emotional response to music (Gregory, 1996). For example, it is possible that our responses could be swayed by the emotion that we believe is ‘culturally’ expected of us. In so doing we line up our own emotional reactions with how we think we should behave in response to the reaction of others (Juslin and Sloboda, 2001). This presents the idea that the emotional state of another person can inadvertently influence our emotions. For example, we might experience joyfulness at a pop concert, as this is the general and expected response (Juslin and Sloboda, 2001). The idea of an expected response becomes particularly apparent on hearing a famous artist or piece of music that we believe to be ‘amazing’ because our culture or

society reflects that opinion (Juslin and Sloboda, 2001). It follows that it is important that the participants of this study do not have any pre-conceived ideas about the vocalist prior to taking part, including any feeling of obligation to respond in a particular way. This will influence the choice of vocalist when structuring the study.

Cultural expectations dictate that we feel *something* when listening to music (Juslin and Sloboda, 2001). Our perception of what emotions we should feel also depends on the type of cultural upbringing we have experienced, “In Western musical tradition the major mode is frequently associated with happiness and the minor mode with sadness” (Gregory, 1996, p.47). This strongly suggests emotional responses to different voice qualities might be influenced by premeditated cultural associations. However, it is also possible that we might not experience any emotion at all when listening to music,“…some people claim that they rarely, if ever, experience emotional reactions to music” (Juslin and Sloboda, 2001, p.97). This argues that the phenomenon of music and emotion may not exist in all people and indeed this could be the case for some of the participants in this study.

Similarly, we could find ourselves experiencing emotion based on our social upbringing. “One example of the strong social influence is the kind of emotion that arises from what one has been told, and therefore may come to believe, about the music in question” (Juslin and Sloboda, 2001,

p.97). The listener might hold specific beliefs about a particular style of music or sound and that could have an impact on the emotional response to it. They might respond with what they perceive to be the ‘correct’ emotion or opinion. This is not an impulsive response yet it is one that is highly influenced by external factors including taste.

When we listen to a song we will often identify and empathise with the vocalist, indirectly reflecting the emotion that they are trying to convey (Jones, 2005). Although we might be able to describe these emotions, it is also possible that they will not be the emotions that we are directly feeling ourselves. Rather we are displaying an understanding of what we believe to be the intended emotional response. This issue was touched upon earlier on in the chapter, during the consideration of emotional contagion. Our response to musical stimuli might be effected by both our personal desires and memories (episodic memory) (Juslin and Sloboda, 2001). For example, if the listener has a memory of their mother singing a lullaby to them as a child in falsetto quality, then the association between this memory and the voice quality used might lead them to be filled with contentment every time they hear said quality. However, if the listener’s mother passes away and the same memory exists, this association could alter the response to different emotions such as sadness and grief.

### *2.4 The difficulties with perception*

Krieman’s study (1993) focusses on the ability of clinicians to rate the patients’ voices on a scale of roughness, this ultimately relies on their

individual perception. In contrast to this, the term perception in this thesis refers to the participants' perception of the different voice qualities as well as their perception of the intensity of the emotional response to the stimulus. This is to say that the participants may feel strongly affected by the stimulus internally but may not respond in the same way externally. This suggests then that it might be challenging to collate authentic emotional responses from the participants.

A potential method for recording emotional responses to voice qualities could take the form of a perceptive based study. The following research acknowledges the significant role of perception on emotional response to stimuli and highlights the advantages and potential challenges posed by perceptive based studies. There are existing standard measurement methods that are currently used to measure the affect and emotion of participants, for example, PANAS (Crawford and Henry, 2004), is utilised to document the patients' emotions, similarly, SUDS (Tanner, 2011) is used to measure the distress levels of a patient. Despite the advantage of using such methods, they are designed to establish the emotional state of the individual, which works well with in a therapeutic setting. However, the focus of this research is to establish whether different voice qualities have the ability to evoke an emotional response with in the listener. Therefore emotional states are not being measured; instead the research is concerned with the affect the stimulus might or might not have on the listener.

As established, individual perceptions of the same thing inevitably vary. Juslin and Sloboda state, "….everyone knows what an emotion is, until asked to give a definition" (2001, p.73). If different voice qualities prove to evoke a range of emotional responses within listeners, then there is a potential issue with the predictability of their responses which arises as a result of their ability to accurately articulate and describe what it is they are feeling. This is because perception, as noted, can be impacted by a number of elements including upbringing, social factors and cultural influences. This is to say that 'feeling content' might mean something different from person to person. How can we accurately measure emotional response to musical stimuli when it is so heavily influenced by individual perception?

The difficulty in successfully acquiring reliable results appears to be a consistent element in perception based studies. That said, some methods prove to be more dependable than others. For example, Juslin and Sloboda (2001) state that collating listener' verbal responses to stimuli presents a useful approach to perceptive based studies. Nevertheless, the flaws surrounding the use of individual perception as a diagnostic tool is highlighted in a clinical setting, in the diagnosis of disordered voices. De Bodt states that there is yet to be a universally designed format for assisting perceptive based studies with regards to voice quality (1997). As a result, speech therapists are required to rely on their own perception of

what is a ‘healthy’ or ‘normal’ voice in order to identify potential trauma or abnormality. Similarly, the use of the patient’s perception is a common way of detecting vocal problems in this setting. It is likely that a patient will seek help if they perceive that there is an abnormality with their voice and it is their perception, once again, that helps to determine whether or not any treatment has been successful. Speech clinicians often use perceptual measures as opposed to absolutes (Krieman, 1993). However, depending solely on perception can be unreliable. This is once again because a ‘croaky’ or a ‘harsh’ sound, for example, could potentially mean something different depending on who is listening and diagnosing (Menasha, 1991).

Voice quality can be described differently depending on the profession or the experience of the listener. De Bodt goes on to say, “To date, research concerning factors influencing perceptual ratings of voice quality has focused on training scales, listening experience, cultural background, and speech fragments” (1997, p.75). Those are some of the existing structures in place for measuring perception. Even so, a definitive way of doing so does not currently exist and perhaps cannot exist. Existing studies based on vocal health are one of the most suitable references to use with a view of developing a study to assist answering the question posed at the beginning of this thesis. Despite the issues surrounding perception, the success of perceptive based studies in a clinical setting suggests that this could be an appropriate method to use for this study.

As mentioned earlier, a reliable system that accurately measures perception has yet to be established. As a result, this presents variables that may be challenging to control. Currently there is not any research that explores the relationship between voice quality and emotion. Notwithstanding, the rest of the chapter provides some insight into previous perceptive based studies from the fields of vocal health and music therapy.

There are many variables to consider when devising a study, as previous studies have suffered from using the wrong form of methodology (Krieman 1993). This accentuates the importance of selecting the most viable methodology for this study. Previous research in similar fields highlights common variables that arise from perceptive based studies. These include the type of candidate taking part and the implications of using a stimulus that is either live or recorded (Jones, 2005). Juslin and Sloboda state, "The reaction may depend on the person, their musical preferences and the location" (2001, p.96). In reference to Tyson's study (Tyson, 2002) the necessity to conduct the study in the same location and at a similar time for each subject is reinforced (2002). This appears to be a consistent finding within these studies, implemented with the aim of controlling potential variables. Gender and location are variables that can be controlled. However, if we choose to eliminate just men from the study or vice versa, we ultimately change the agenda of the study to focus on how

voice quality may arouse a response within one sex as oppose to within listeners in general. Regarding location, this can be controlled by ensuring that each participant takes part in the study at the same time and in the same place. In summary, cutting out an entire group would inevitably limit findings and prevent complete exploration of the question in hand. Nevertheless, it is important to note that the musical preferences of each individual cannot be controlled and this could influence their response to the stimulus.

As mentioned earlier, our musical preferences can be influenced by cultural and social factors, as well as our individual up-bringing and exposure to different styles of music. For example, if one is familiar with classical singing and unfamiliar with distortion (refer to methodology chapter for information on distortion), the latter may be perceived as ugly, uncomfortable or even dangerous. This perception would inevitably affect the emotional response that follows and would differ depending on the taste of the individual. This leads to another variable for consideration known as internal standards. This is referred to in Krieman's paper (Krieman, 1993), which is based on the reliability of the perception of different voice specialists in the diagnosis of disordered voices. Krieman found that we use our own internal standards to describe the sound of a person's voice in relation to descriptors such as 'roughness' (1993). As different degrees of 'roughness' are likely to mean different things from person to person, there is no absolute way of defining the overall condition

of the voice in question utilising perception alone. Krieman (1993) asked raters to judge the roughness/ hoarseness of subjects' voices by listening to them. Significantly, the raters were professionals with knowledge in this area; however, they are also human beings. Therefore, it is fair to assume that alongside their specialised knowledge, they will have used their internal standards to support their judgments. The study that will be devised to assist answering the question posed at the beginning of this thesis will not ask participants to judge the quality of the singer's voice, but instead to express their emotional response to what they are hearing.

Jones devised a study (Jones, 2005) whereby chemically dependent subjects took part in various types of music therapy sessions, with the aim of discovering which were most effective. Results showed the possibility that pre-recorded and live performances could elicit different responses (Jones, 2005). Whether the stimulus is presented as live or recorded may give rise to further variables. Similarly, Bunt's study of music and emotion using improvisational music therapy found that live performances highlighted further variables. These included the body language, movements and facial expressions of the musicians (Bunt and Pavlicevic, 2001). This suggests that being able to see somebody deliver a performance as opposed to simply hearing them could have an impact on the listeners' responses. For example, if the performer happened to cry during the performance, the audience might empathise with them and consequently feel the urge to cry themselves. This is to say that live music

could potentially complicate results by creating additional visual variables and preventing the participants from responding to the music alone.

Tyson's perceptive based study (Tyson, 2002), consisted of a small group of subjects in a hip hop therapy intervention, who were asked to express how the therapy made them feel. The study proved that the therapy was not very successful and this may be due to the fact that only a small set of results were obtained from a small group of participants (Tyson, 2002). With this in mind, additional questions arise surrounding the optimal number of participants for this study. A small group could narrow results, however, a larger group has the potential to present further variables that might hinder the success of the study. An important feature of Tyson's study (Tyson, 2002) was the inclusion of both a pre-test and a post-test (Christensen et al., 2014). A pre-test serves to determine each individual's walk of life prior to the study taking place. It could well be useful to gain some understanding of an individual's perception of different voice qualities and musical styles as dictated by experience, exposure and up-bringing. The reason being that the demography of the participants could potentially have an effect on the results of the study. For example, if half of the participants are vocalists and the other half are not it is possible that the vocalist's responses might be influenced by their knowledge, therefore influencing the overall results. This information could be useful when gathering potential reasons that exist behind particular responses. Background information could make analysis more in-depth as well as

lessening some of the problems surrounding perceptive based answers. On the other hand, the use of a pre-test might unintentionally change the focus of the study. Analysis of results in conjunction with individual background information would be based on assumption. Therefore, a pre-test might complicate and cloud the results. A post-test could potentially occur as a repeat of the pre-test, in order to determine whether the participants' answers have changed as result of their initial experience of this study. However, such testing is not relevant to this study as it will focus on discovering whether the participants experience an emotional response to the different voice qualities used, as opposed to the ability of the experience to change the participants.

In addition to the variables mentioned above, audio processing would also need to be considered. For example, the use of compression and reverb creates a more refined and atheistically pleasing sound, which could potentially have an affect on the participant's responses. Similarly, absolute pitch, the various dynamics of the different voice qualities and the order of which the voice qualities appear might also impact the participants' responses. Therefore these variables need to be considered when devising the study.

The information outlined above provides an insight into the ways in which typical variables can be minimised or avoided. The methodologies utilised in the following studies have been outlined to assist the development of an

original study that aims to answer the question posed at the beginning of this thesis.

In Gregory's study (1996), he provided listeners with a list of adjectives to utilise when recording their responses. He asked them to listen to a variety of Western classical, Indian classical and New Age music (Gregory, 1996). They were instructed to select the adjectives from the list provided that best conveyed the mood of the piece. This list was modified from Hevner's adjective list (Hevner, 1936), and presented a structured and reliable way for the listeners to express how the stimulus made them feel. Similarly, Juslin and Sloboda describe some ways of measuring emotional responses by using adjective lists, scales and ways of expressing personal responses (2001). Ultimately, these methods enable the listener freedom to express how the stimulus made them feel, yet with a structure such as a scale to refer to. Another example can be found in a study conducted by Jones (2005), whereby he used a scale between zero and eleven. Zero represented no response and eleven an intense response (Jones, 2005). These results were then used to discover how much the listener was affected. Jones went on to say, "The constructed scale allowed for changes in emotions without requiring a correlation between positive and negative emotions" (2005, p.100). Jones's adapted scale offers another way for listeners to rate their responses without imposing or assuming anything. By using numbers as opposed to adjectives, one can measure the intensity of the emotion, regardless of what that emotion is.

Opting to not define specific emotions eradicates some of the issues surrounding the reliability of individual perception. This was something that was considered when structuring the study featuring in this thesis.

Krieman (1993) was unable to decipher the best method to use when measuring voice quality amongst disordered voices. This implies that measuring perception is a difficult task regardless of which method is used. The issues that have arisen from existing studies lead to further questions worth exploring within this study. Is it possible to accurately measure perception of both voice quality and emotion? Is there a way in which we can standardise perception of voice quality? Would this mean exposing the participants to primary colour audio examples of the different qualities and presenting them as absolutes? Can we standardise perception of emotion? How can we ensure that participants are able to articulate themselves reliably? This study might bring us closer to answering some of these questions.

Other methods for establishing listeners' responses to musical stimuli include scientific and physiological approaches, "…researchers of emotion have used a number of physiological indices to measure emotion, including heart rate, respiration, skin conductance, muscle tension, electrocardiogram (EKG), blood pressure, and electroencephalograph (EEG)" (Juslin, 2001, p.74). This would present a more definitive way of establishing an emotional response, however may not benefit this

particular study. This is because the controlled environment in which the study is likely to take place already adds a sense of formality. Adding machinery would enhance this, removing the participant further away from a relaxed, comfortable setting whereby they would most likely experience a spontaneous response to what they are hearing. For this reason, a perceptive based study proves to be a more appropriate approach than a scientific experiment. The question remains as how best to enable the group of listeners participating in this study to record their emotional responses to the voice qualities they are hearing. In particular, how might the role of unconscious factors be accommodated within this research? These questions will be explored in the methodology section where the work of D.W Winicott, and in particular the role of transitional objects as mediators between unconscious factors and consciousness, will beincorporated into the research method of this study.

# 3. Methodology design

Substantial research has already been conducted connecting emotional response to musical stimuli. Additionally, literature based on perception of voice quality in the assessment of disordered voices already exists. However, there is a lack of research focusing on emotion evoked solely through the use of different voice qualities in the context of singing. Equally the ecological approach adopted by Eric Clarke (2005) has the benefit of being wide ranging: by contrast this research question is specific and focused. Clarke (2005) identifies a range of elements such as, psychological and musicological perspectives as well as historical, social and cultural factors. Conversely this research concentrates on different voice qualities as a means to elicit an emotional response from the listener. These differing approaches highlight the need for further research within this field. As a result, an original study was devised in order to assist answering the question posed at the beginning of this thesis. This chapter will give a chronological description of the development of the study, including the different methodologies that were considered, reasons for rejecting them and justifications for the approach that was finally chosen.

Further information surrounding Jo Estill's perceptual study can be found in *Estill Voice Training System Level Two: Figure Combinations for Six Voice Qualities.* As mentioned in the plan of thesis section, part of Jo Estill's extensive research involved conducting a perceptual study,

whereby participants listened to four modes and were asked to identify them. The study was successful in that all of the qualities were identified accurately, with the exception of opera. As a result of acoustic analysis, it was discovered that, “The acoustic characteristics of opera quality combined characteristics of Speech, Sob, and Twang” (Klimek, et al., 2005b, p.3). The fact that opera quality is essentially a mixture of varying degrees of three other modes suggests why it was challenging to identify. The success of Jo Estill’s perceptual study on voice qualities suggests that a perceptive based method could be the best way of conducting this original study, which also focuses on voice qualities. Conversely, the inability of participants to accurately identify opera quality in Estill’s study highlights the potential issues that could arise from responses heavily reliant on individual perception.

A unique composition provided the basis for the original study. A chord progression was produced on the piano to provide inspiration for the vocalist and recorded into Logic Pro X. After listening to the accompaniment, the vocalist was able to devise a concept that provided the stimulus for the lyrical content. The words of the song conveyed an ambiguous description in the third person about a soul trapped inside the shell of a body that no longer functions due to terminal illness. As the composition progressed to the chorus, the lyrics described how the lady had passed away and was now free from suffering. Below are the lyrics as they appear in the song:

***Verse 1***

*All alone, inside four cold walls,*

*She's lost with just old memories.*

*Emotional, she has so much to say,*

*Powers bound, her freedom's taken away.*

***Chorus***

*But it's ok now, you can smile again,*

*The clouds have made a way for you.*

*It's ok now, you can laugh again,*

*The stars have made a place for you.*

*Now the angels will hold you x2*

***Verse 2***

*Staring out, wishing to be heard,*

*Stories being told through her eyes.*

*Expressionless, she has so much to say,*

*Powers bound, her freedom's taken away.*

***Chorus***

*But it's ok now, you can smile again,*

*The clouds have made a way for you.*

*It’s ok now, you can laugh again,*

*The stars have made a place for you.*

*Now the angels will hold you x2*

***Bridge***

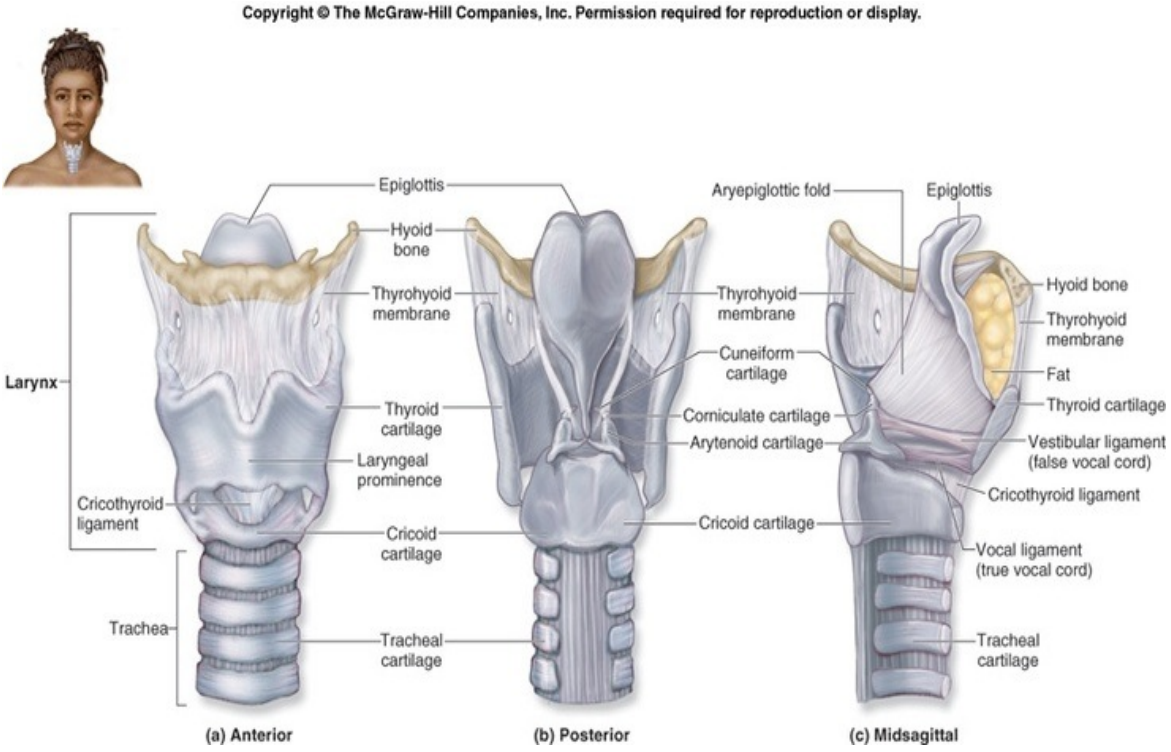
*Yeah, and the angels will hold you (ad.libs)*

(Refer to track 1 to hear the original composition)

During the process of writing the lyrics, a melody was also created. The vocalist featured on the composition had prior understanding and experience of producing the different voice qualities outlined by Jo Estill. With the lyrics serving as a stimulus, the vocalist was free to respond to the intention of the words through the use of these different qualities. The vocals were then recorded over the accompaniment; using Logic Pro X.

Figure 1 is an image of the different cartilages, muscles and structures that make up the larynx, followed by a description of the production of some of the voice qualities defined by Jo Estill along with their permutations. Perceptual acoustic correlates and emotions commonly associated with these qualities are also mentioned.

**Figure 1- The structures that make up the larynx**




(The Mcgraw-Hill companies, Inc (2013).

(Refer to the accompanying audio files to hear examples of each of the following qualities.)

**(Track 2)**

Speech quality requires that the true vocal folds close thickly, providing a strong resistance to breath. In order to produce thick fold contact at a low muscular effort level, the larynx must be in a vertical, relaxed position. When the true vocal folds (see appendix A for information on true vocal folds) are thick, there is some contraction within the body of the folds. The cover is loose, providing mucosal wave (a ripple from the bottom to the top

of the vocal fold edge) (Klimek, et al., 2005a). Mucosal wave provides harmonics, therefore, thick vocal fold mass results in a sound with some core and presence. All other structures remain in their mid-relaxed states (Klimek, et al., 2005b). As well as in everyday speech, speech quality can be heard across a range of musical genres including pop, folk and musical theatre (Beechler et al., 1996). Speech quality makes diction highly comprehensible and the words can be heard with ease, reinforcing the lyrical content. Although there might not be a specific emotion associated with this quality, audiences can relate well to it as it is close to every-day speech and not masked by other things commonly added by the vocalist.

**(Track 3)**

Cry quality is produced when the thyroid cartilage tilts forward as a result of the contraction of the cricothyroid muscle and the true vocal folds are elongated and stretched as a result. “The Thyroarytenoid (TA) and cricothyroid (CT) are the intrinsic muscles of the larynx that regulate the length of the true vocal folds” (Klimek, et al., 2005a p.21). The stretching of the true vocal folds is integral in order to access higher pitches with ease. This is because the true vocal folds have to vibrate more times per second higher in the range and cannot do this easily if they remain too thick (Klimek, et al., 2005a). In addition, tilting the thyroid cartilage forward allows the singer to access the cry posture/quality which can be likened to that of a whining child (Klimek, et al., 2005b). This same maneuver occurs when we cry in real-life. Unknown to a listener, on hearing this quality they

might mimic the laryngeal posture of the vocalist which could trigger feelings of sadness and perhaps even tears as a result. Cry quality is often used to enhance a deep emotion such as passion (Kayes, 2004). The act of tilting the thyroid cartilage forward might be the difference between a spoken and a sung sound.

**(Track 4)**

Jo Estill defined sob quality as being the suppressed sobbing of an adult that mourns (Klimek, et al., 2005b). To access the sob posture, the thyroid cartilage is tilted at a high effort level and the larynx must be held in a very low position (Klimek, et al., 2005b). In order to maintain this extreme and unnaturally low position, the extrinsic muscles of the head and neck and torso are contracted so as to stabilise the larynx (Klimek, et al., 2005b). Sob quality is perceived as a dark, soft tone. It can be heard in operatic pianissimo, as well as in Blues and Jazz and offers an emotionally intense sound (Klimek, et al., 2005b). The lowered laryngeal position and the high muscular effort required to sustain it, replicated in the listener, can cause them to sob for real, triggering feelings of grief and despair.

(**Track 5)**

The defining feature of falsetto quality is the raised vocal fold plane. This occurs when the arytenoid cartilages rock back and out, raising the back part of the true vocal folds and pulling them apart (Klimek, et al., 2005b). The body of the true vocal folds is stiff, resulting in few harmonics above

the first formant (Klimek, et al., 2005b). All surrounding structures are in their mid, relaxed states as in speech quality. (Klimek, et al., 2005b). With falsetto quality there is a short closed phase, meaning that the true vocal folds remain closed for a short period of time each vibratory cycle (Menasha,1991). There is minimal obstruction to the airflow with very little if any vocal fold contact, which is why falsetto quality may be perceived as breathy in nature. This quality has a child-like essence about it and it can be *"…useful for moments of vulnerability, intimacy and uncertainty"* (Kayes, 2004, p.154).

**(Track 6)**

The primary structural feature of belt quality is cricoid tilt. When the cricoid cartilage tilts forward the true vocal folds bunch up and become as thick as they possibly can be (Klimek, et al., 2005b). As a result, the vocal folds remain closed for a large proportion of each vibratory cycle. This provides an increase in subglottal air pressure and an intense sound pressure wave (a very loud sound) (Kilmek, et al., 2005b). The Aryepiglottic Sphincter narrows providing a formant between 2-4 kHz (see appendix A for information on the Aryepiglottic Sphincter). This is the area in which the ear hears best, therefore the sound is perceived loudly in the ear of the listener (Klimek, et al., 2005a). These components are what make belt quality such an intense, exciting sound. Belt is an everyday quality commonly heard as a yell or a cheer (Estill, 1988), as well as being the first sound produced by a new born baby (Klimek, et al., 2005b). The

intensity of belt quality is likely to grab the attention of the listener. It *"…can be used to express a range of heightened feelings: joy, despair, anger, frustration, exuberance…"* (Kayes, 2004, p.156).

**(Track 7)**

Control of the false vocal folds is outlined in the Estill Model (see appendix A for information on false vocal folds). Constriction of the Ventricular folds, as they are otherwise known, can contribute to the production of a distorted sound (Obert and Chicurel, 2005). However, Jo Estill did not define false vocal fold constriction as a quality itself. If the false vocal folds begin to move in toward one another during phonation, they impede on the free vibration of the true vocal folds (Klimek, et al., 2005a). This results in a tight, pressed sound (Klimek, et al., 2005a). It has not yet been established whether it is possible to constrict the false vocal folds without compromising the free movement of the true vocal folds (Klimek, et al., 2005a). Consequently, use of this sound can result in vocal trauma. Distortion can create an exciting sound but may also be perceived as angst or aggression.

**(Track 8)**

Another form of distortion is described in the Estill Model as "… *glottal fry, or creaky voice."* (Klimek, et al., 2005a p.44). To produce this sound, the true vocal folds are slack. Both the body and the cover are loose resulting in an irregular vibratory pattern (Klimek, et al., 2005a). There is no pitch

associated with slack folds. To access glottal fry, the larynx must be vertical and relaxed[2]. It can be used to give an sense of relaxation and has a casualness about it. Glottal fry leads nicely as an onset into speech quality.

The qualities outlined above have been described in a primary colour fashion. In the context of singing, one is more likely to use combinations of these, or indeed a permutation of the original recipe.

**(Track 9)**

The addition of thyroid tilt (cry quality) is essential in order to take any voice quality higher in the range. (Klimek, et al., 2005a). Speech quality is no exception. In order to keep speech quality consistent as you ascend the range, a degree of thyroid tilt is required to access the pitch safely. However, it is important that the vocal folds are not elongated too much, as the mass of the folds must remain fairly thick in order to maintain a speech-like sound (Speed, 2010). This permutation is known as speech and cry (Speed, 2010). In addition, it is the presence of thyroid tilt that allows us to access a true vibrato and to sound like we are singing rather than speaking (Klimek, et al, 2005a).

[2] P.Penny, 2015

**(Track 10)**

Adding thyroid tilt to falsetto quality can refine the sound, *"One can sweeten the tone by titling the thyroid cartilage slightly…"* (Klimek, et al., 2005b p.25). However, the thyroid cartilage must not tilt too much when in falsetto quality as this will increase vocal fold contact, resulting in a less breathy sound. This combination could be described as falsetto and cry.

**(Track 11)**

A breathy speech sound is the result of incomplete vocal fold closure (posterior glottal chink) on a horizontal plane[3]. The gap between the posterior portion of the true vocal folds provides less resistance to the breath, allowing it to escape and resulting in a breathy tone (Chicurel & Obert, 2005; (Klimek, et al., 2005b; Speed, 2010). However, unlike falsetto quality, there is still core and depth to the sound.

**(Track 12)**

When the Larynx is lowered, the lower harmonics are amplified providing depth, darkness and bass resonance (Klimek, et al., 2005a). When this occurs with thick vocal fold mass and some degree of thyroid tilt, one can access a louder and less emotionally intense variation of sob quality/a darker version of speech quality[4]. One may refer to this permutation as speech and cry with a lowered larynx.

[3] P.Penny, 2015
[4] P.Penny, 2015

The lyrical content served as the stimulus to inspire the vocalist's choice of voice qualities. The author of this thesis was the chosen vocalist and an experienced singer. As a result of having a highly flexible and responsive larynx, one may argue that the choices surrounding voice quality were made impulsively. Nevertheless, for the purpose of the current study, it was important that a variety of vocal qualities were included within the performance of the song, therefore, some of these choices were undoubtedly made consciously. Table 1 includes a description of the voice qualities (as defined by Jo Estill) used by the singer at different points in the composition (Refer to Appendix C for the complete table).

**Table 1- Description of voice qualities**

| Section | Lyrics | Voice quality |
|---|---|---|
| **Verse 1** | All a- | Breathy speech |
| | -lone inside four cold walls | Falsetto and cry |
| | She's lost with just old memories | Breathy speech |
| | E- | Glottal fry/creak |
| | -motion- | Speech and cry |
| | -al she has so much to | Falsetto and cry |
| | say | Breathy speech |
| | Po- | Speech and cry with lowered larynx |
| | -wers bound her freedoms taken away | Breathy speech |
| **Bridge 1** | But it's okay now you can smile again | Speech quality with minimal thyroid tilt/thick folds |
| | The | Glottal fry/creak |

| Section | Lyrics | Voice quality |
|---|---|---|
| | Clouds have made a | Speech and cry with lowered larynx |
| | Way | Glottal fry/creak |
| | For you | Speech and cry with lowered larynx |
| | It's okay now | Speech and cry |
| | You | Falsetto and cry |
| | Can laugh again | Speech and cry with larynx in a mid-position |

Once the composition was complete, with the research question in mind, the next step was to design a study that measured emotional response to the different voice qualities featured in the song. In order to control the harmonic accompaniment as a variable, it was removed. Similarly, any intention to include harmonies and backing vocals was abandoned. This was an attempt to strip away all elements, other than the voice, that might have an impact on emotional response. Lyrical content remained as a variable; this decision will be discussed in more detail further on in the chapter.

(Refer to track 13 for a recording of the pilot study)

With the composition providing the stimulus, a pilot study was devised that aimed to measure emotional response to different voice qualities in the form of an interview. Once completed, the results of the pilot study were

analysed in order to distinguish what was successful and what could be improved. Initially, the interviewer asked one participant a series of questions about their responses to the composition. These questions were designed to trigger discussion and to allow the participant to answer in a detailed manner. The composition was played twice through from beginning to end. The interview questions that were posed are outlined below.

Interviewer: *How do you feel after listening to the composition? Has the vocal moved you in any way?*

Interviewer: *What emotions do you feel the vocalist is trying to evoke through their choices of voice quality?*

Interviewer: I am going to play the composition again. *Please could you raise your hand at any point that you feel the vocalist is conveying a particular emotion or where you feel there is a distinct change in the emotion conveyed.* I will then stop the music and you will have a chance to expand verbally on what you have experienced.

Interviewer: *Which factor do you feel was most influential in moving you emotionally: voice quality, lyrical content or an amalgamation of the two?*

On analysis of the results, the questions proved to be too broad and vague. Although an emphasis had been put on avoiding a controlled interview, it might have been more successful if the questions had enabled the listener to discuss something specific. The pilot study was conducted with a participant who was a Master of the Estill Model. It became apparent throughout the interview that her answers were based on what she believed to be fact rather than a spontaneous emotional response. The participant's answers were highly technical, more so than your everyday average candidate. She was inevitably highly influenced by her knowledge and experience. As a result, this posed further questions about the type of participants to select for the study. Additionally, this observation further reinforces the fact that the style of questions used was inappropriate.

The issue of recording emotional response to voice quality proved to be challenging in the pilot study. This difficulty suggests that there may be benefits to recording the listener's responses in real time, as opposed to asking them to describe their responses after they had already occurred. Listening to a song all the way through before responding could cause the listener to forget what it is that they felt at the time. Similarly, the act of pausing the music in the second part of the pilot study, allowed for the participant to be distracted and to consider their response rather than reacting spontaneously. These elements give rise to other potential variables. It would be more reliable to capture the listener's responses in

the moment. This also takes into account those participants that don't have any formal musical training and as a result, may not be able to easily articulate what it is they are experiencing. A simpler way of recording listeners' responses would provide much more specific results than the previously broad, open questions that featured in the pilot. Additionally, this would assist the analysis process and ultimately help to control some of the variables.

Other questions were raised as a result of the pilot study such as, how much of the listener's emotional responses were based on the voice qualities and how much were they influenced by the lyrics and/or construction of the melody? The fact that the composition presented a solo voice eliminated other musical variables, but did not eradicate the potential impact of the melodic contour or lyrical content. Could the words be removed and the melody performed on vowel sounds instead? This may however lead to issues regarding the ability of the vocalist featured in the composition to interpret the piece artistically without the presence of lyrics. If it is the lyric that dictates thought, intention and subsequent choice of voice quality, how can the vocalist deliver a performance if their stimulus, in this case, the lyrics is removed? In order to achieve artistry and to communicate well, voice quality ought to be text driven. Moreover, when the larynx is flexible, the voice will respond instinctively to the thought/intention of the lyric. Removing the lyrics would essentially remove the impulse for the vocalist. This could mean that it is impossible to

discover whether voice quality alone evokes emotion. If a set of lyrics are sung in one consistent voice quality without any variation, they would become meaningless from a communicative point of view. If a melody was sung with interpretation and feeling minus the lyrics, the message/content would not be specific. This is to say that the listener may perceive elements of sadness in the sound, for example, but that is quite a generalised response. Is it then the lyrics that enable the vocalist to be specific in what they are communicating? Perhaps it is an amalgamation of the words, melody and voice quality chosen to reinforce them? Similarly, the melodic contour itself also presents another variable. How can one be sure that it is the voice quality chosen that evokes an emotional response within the listener and not the construction of the melody? Pitch and rhythmic characteristics of the tune inevitably have the ability to move a listener. It is possible to remove the melody altogether and have the vocalist perform the lyrics in the speaking voice as the voice qualities used are not exclusive to singing. However, it is impossible to eradicate both the lyrical content and melody as variables without removing the stimulus altogether. These elements ultimately make the original question harder to answer and require further consideration.

### *3.1 Implementing the methodology*

The method used in the pilot study was unsuccessful, therefore, research continued with the view of revising and changing the methodology. Even so, the composition remained as the stimulus for the study. Qualitative

research proved usual, as it allowed a focus on how the participant in question feels and thinks about the stimulus. Drawing on other studies in this field of research, this thesis used both qualitative and quantitative methods. This is because the study is attempting to establish if different voice qualities evoke an emotional response and if so, to measure the intensity of the responses. Therefore, this study will rely on the participants' expression of emotion in order to attempt to draw a conclusion.

With the composition the only remaining component of the pilot study, a new method needed to be designed to better serve the research question. Background research collated from various pieces of literature gave some insight into how to approach perceptive based studies. After examining previous studies from similar fields of research, the use of a scale or adjective list to record emotional response to the stimulus appeared to be the best approach for this study. However, this was later rejected due to the fact that the participants' responses could not be monitored in real time. The pilot study highlighted the importance of documenting the listeners' responses in the moment. Waiting until the end of the composition would allow the participants to think consciously rather than spontaneously about their reactions. Additionally, they might not be able to recall how they felt if too much time had passed. The pilot study also flagged up questions about what was being measured. The questions devised for the interview placed an emphasis on the *type* of emotion being

experienced as a result of different voice qualities. As highlighted in the literature review, emotion can be difficult to describe and emotional experiences differ from person to person. Consequently, it was decided that this study would measure the intensity of emotions experienced as opposed to the type, with the view of analysing the specific voice qualities that evoked the biggest emotional reactions.

Intensity of an emotional response within the context of this thesis refers to the degree that participants have been affected by the stimulus. A similar approach was utilised by Jones' study (2005) in which he used a scale to determine the intensity of the participants' responses to different types of therapy. Jones was documenting an emotional response focusing on its intensity, which is also the concern of this study. The two dimensional arousal model for the discrimination of a range of emotional intensities plots different emotions according to their strength i.e. anger, terror and fear. Such a discriminatory approach indicates one area for future research, however for the purposes of this study it was decided to focus on intensity (in the two-dimension model as referred to as arousal) only. The use of an intensity scale was considered. However, as mentioned earlier, it was ultimately rejected on the basis that it did not enable the responses to be recorded in real time. With neither an adjective list nor scale proving to be an appropriate method, research continued to find the approach that best served the research question. The study needed to be conducted in a way that enabled the participants to respond

spontaneously. Without wiring participants up to machinery, there wasn't a way of avoiding a perceptive based study. This option was considered but decided against due to concerns that it would be too invasive. If the participants were asked to articulate their responses, this would more than likely instigate conscious thought. The initial reaction would have been and gone before it could be explained. Consequently, authentic responses may be missed by attempting to record results in this way. It appeared that monitoring responses in real time would be impossible if participants were asked to verbalise their feelings. Donald Winnicott's (1958) concept of transitional objects helps to explain this space between the experience of an emotion and being able to make conscious sense of it. Transitional objects refer to a stage in the development of infants but they are also relevant to adulthood. They are based on the idea that there is 'something' between us and the rest of the world that is not documented. When we hear a piece of music, we may feel something, the transitional object is the space between experiencing that feeling and responding to it.

Associations made between past experiences and specific voice qualities may serve as transitional objects. This was addressed in the Literature review in reference to the evaluative conditioning mechanism of the BRECVEM model. In a similar manner, Jo Estill states that observers of her perceptual study "…make their judgments within the context of some kind of psychological space" (Colton and Estill,1981,p.334). She likens this to the mathematical space known as the 'Euclidean Space' (Colton and Estill, 1981). Estill notes that although subjects could accurately identify

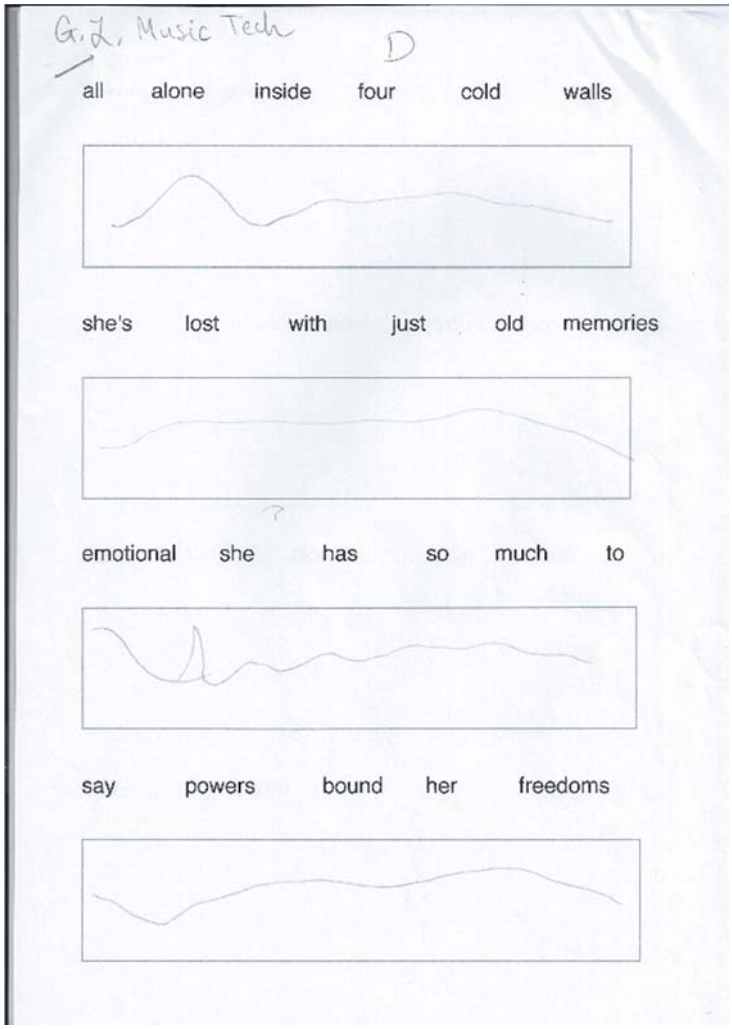
different voice qualities, it is not possible to ascertain how they arrived at such judgments (Colton and Estill, 1981). This further reinforces that what a particular emotion means to one person, can mean something completely different to another person. If a participant expresses that the stimulus made them feel depressed, how do we know that their version depressed is the same as our perception of depressed? This supports the idea that measuring intensity of emotional response to voice quality would be more feasible than attempting to record the different types of emotions experienced.

Donald Winnicott was a pediatrician and psychoanalyst. As part of his clinical practice he devised the squiggle game to enable young patients to express their emotions through drawing (Chescheir, 1985). The game begins with the therapist drawing a squiggle on a piece of paper and asking the patient to transform it into a picture. The idea is that the picture that develops represents subconscious feelings that may not otherwise be verbalised. The therapist can then use the image to gently probe for further insights into the emotional state of the individual. Although the squiggle instigates conversation between the patient and therapist, the image begins as something completely random that then develops to unravel subconscious feelings as demonstrated in the video by Enfield council (EnfieldCouncil, 2014).

The spontaneity and non-invasive nature of the squiggle game (Chescheir, 1985) inspired the chosen methodology for this study. It was modified to enable participants to respond to the stimulus independently and without the need for conversation. It was also taken into account that those taking part in the study would be adults not children. As adults tend to be more self-aware than children, simply being asked to draw a picture may invite all kinds of conscious consideration. Therefore, instead of drawing an image to express their emotional responses, it was decided that participants would draw a line as they listen to the composition, enabling the hand to move freely to display the peaks and troughs of their emotional responses. Ultimately, this method was chosen because it proved to be the most impulsive way of recording intensity of emotional response, without relying on verbal description and explanation from the participants. This method also enables the participants' responses to be recorded in real time. The participants were given a booklet including the lyrical content, to serve as a guideline for them to follow as they listened and to scribble their responses. This would serve as useful in the analysis process, in terms of distinguishing which voice qualities evoked the most intense emotional responses. See figure 2 for an example of one of the participant's booklet.

(Refer to appendix B for each participant's booklet).

**Figure 2- Example of participant booklet**




Note the peak in emotional response on the words ‘alone’ and ‘she’ (figure 2). In both places, the vocalist accessed falsetto and cry quality. In addition, there is a gradual climb in emotional response documented from the word ‘power.’ The vocalist used speech and cry with a lowered larynx on this word. Correlations between intensity of emotional response and these qualities are discussed in detail in the analysis chapter of this thesis.

Inter-rater and intra-rater reliability was also considered when developing the study. Inter-rater reliability is the agreement amongst the different participants each time the study is repeated (Gwet, 2014). Intra-rater reliability refers to the consistency of responses from one participant on more than one occasion. Inter-rater reliability isn't relevant to the current study as the question is not concerned with participants giving the 'right' or 'wrong' answer. This is because the study will be perceptive based. Inter-rater reliability appears to be more relevant within a clinical setting, whereby the agreement between medical professionals regarding vocal health of a patient is imperative to their successful treatment and recovery. Initially, repeating the study more than once was considered, placing an emphasis on intra-rater reliability. It is possible that the participant may change his or her response after hearing the stimulus for a second time. Although this may improve reliability with regards to a single person's response, the idea to repeat the study more than once was later rejected on the basis that participants would undoubtedly have developed pre-conceived ideas about the stimulus after already having heard it. On the other hand, one might argue that the analysis process could potentially give rise to intra-rater reliability. For example, the participants' responses might be compared to one another in order to establish any patterns that might lead to answering the research question. Notwithstanding, the decision was made not to include intra-rater reliability in the analysis process.

Initially, the idea to have a pre-test (Tyson, 2002) seemed beneficial in order to establish the type of people participating in the study. However, after much consideration, the pre-test process was eliminated. This was due to the fact that knowing the participants' ages and musical preferences/experiences, unavoidably changes the focus of the study. For example, this may flag up questions such as; do those participants that prefer rock music have a similar response to the stimulus? This kind of analysis would not serve the research question as it is not concerned with unearthing cultural and social reasoning behind participants' responses. In summary, including a pre-test in the case of this study would further complicate and ultimately change the focus of the question in hand.

Prior to conducting the study, another pilot took place with two anonymous participants in order to assure the new method was viable. It was decided that the pilot would go ahead under controlled conditions in a music studio. The participants were seated and in front of them was the booklet containing the lyrics featuring in the composition. The author briefed the participants on what they were expected to do and told not to communicate with one another. It was explained that they should draw a line in the box provided below the lyrics as they listened to the composition and that a squiggle in an upward direction would indicate a high intensity emotional response and a squiggle in a downward direction, a low intensity response. The composition was then played in its entirety with the participants 'squiggling' their responses. The method appeared to be

successful in that some clear responses to the stimulus had been documented. The squiggles displayed some consistencies between the participants. For example, each time the chorus occurred, both participants' illustrations displayed a strong emotional response. In particular, the word 'angel' appeared to elicit an intense reaction from both participants. The different qualities that were used on this word were falsetto, breathy speech and speech and cry. Although both participants exhibited high emotional responses on this particular word, participant one reacted most strongly when this was sung in speech and cry quality, whereas participant two experienced a stronger response when it was sung using falsetto and cry. This highlights the possibility that different voice qualities could evoke an emotional response within the listener; however, as the quality that evoked the strongest reaction was not consistent between participants, this suggests that it could be the chosen word or construction of the melody that instigated such responses. Speech and cry with both thicker and thinner folds caused strong emotional reactions from both participants, as did speech and cry with a lowered larynx. The subconscious replication of both cry and sob posture may be responsible for these responses.

There were also many differences highlighted in the results. The unique ways in which the two individuals squiggled was interesting. One participant drew large wavy lines and the other produced an image that was much more defined and detailed. This implies that some participants'

responses might be easier to decipher than others when it comes to the actual study.  For example, the participant whose peaks and troughs were more intricately illustrated might be easier to analyse due to the detailed squiggle. In contrast, the participant who recorded their responses using large wavy lines produced a more vague documentation that might be more challenging to decipher. It appeared that one participant had a stronger response overall than the other. This participant's responses always began in the middle of the box and always proceeded to go up, whereas the other participant generally began drawing their squiggle from the bottom of the page.

Both participants agreed that the booklet was difficult to follow whilst listening to the composition. They mentioned that they found it hard to keep up at times and that turning the pages disrupted the flow of the study. However, they also explained that this was a small issue and that they were still able to respond to what they were hearing in real time. As a result, the booklet was not revised.

The participants mentioned at the end of the pilot that they definitely felt emotionally aroused during the study. However, they weren't sure if this was a result of the vocalist, the lyrical content or the general mood of the composition. Despite this insight, for the reasons outlined previously in this chapter, neither the lyrical content nor the melody was eradicated. Although the presence of these variables suggested that it was going to be

difficult to ascertain whether emotion can be evoked *exclusively* through the use of different voice qualities, it was clear that a reaction of some description had occurred and this proved worthy of further exploration. With the second pilot study proving a success, the actual study could commence.

Eight people participated in the study which took place in a group setting in order to eradicate any extra pressure being placed on the individual. In a perceptual study conducted by Colton and Estill (1981), of the thirty participants involved, half were experienced in listening to different qualities (a selection of singers, instrumentalist and speech and language pathologists) and the other half were naive (Colton and Estill,1981). All participants were given training in identifying the different voice qualities and had the opportunity to practice doing so before two formal listening sessions occurred (Colton and Estill,1981). Participants had the task of allocating each stimulus to one of four unnamed modes (Colton and Estill, 1981). Estill drew comparisons between the ability of participants with varying levels of experience to accurately identify each quality; therefore, a large group was required.

This study was not concerned with the ability of participants to accurately identify different voice qualities; therefore, a small group of participants was sufficient. There were no pre-requisites for the type of individual taking part in the study. For example, there was no emphasis placed on

how musically trained the listener was, as this would change the focus of the study. The participants were not informed of the research question nor were they told who the vocalist was or that they would be demonstrating a variety of different voice qualities. Instead they were told to maintain contact between the pencil and paper within the box provided and to allow the hand to move freely to express an emotional response to what they were hearing as the song progressed. Participants were informed that a squiggle in an upward direction signified an increase in intensity of emotional response and a squiggle in a downward direction, the opposite. The purpose of implementing the study in this way was to obtain spontaneous results by limiting participant interference as far as was possible.

At the end of the study, the participants were informed that the vocalist featured in the composition was the study representative herself. This was not mentioned prior to the study being conducted in order to prevent participants from feeling obliged to respond in a certain way. The reaction from the participants when they learned that the singer was the study representative confirmed that had they had this information beforehand, it would most certainly have influenced their responses. Additionally, the different voice qualities that were utilised in the song were briefly explained to them in order to put the study into context.

# 4. Analysis of the study

As mentioned in the methodology chapter of this thesis, prior to the study taking place, the voice qualities that were used by the vocalist at each point in the song were analysed and are detailed in table 1. This chart indicates the qualities demonstrated with the corresponding lyrics. Table 1 shows a section of the chart that appears in the methodology chapter.

(Refer to Appendix C for the complete table.)

| Section | Lyrics | Voice quality |
|---|---|---|
| **Verse 1** | All a- | Breathy speech |
| | -lone inside four cold walls | Falsetto and cry |
| | She's lost with just old memories | Breathy speech |
| | E- | Glottal fry/creak |
| | -motion- | Speech and cry |
| | -al she has so much to | Falsetto and cry |
| | say | Breathy speech |
| | Po- | Speech and cry with lowered larynx |
| | -wers bound her freedoms taken away | Breathy speech |
| **Bridge 1** | But it's okay now you can smile again | Speech quality with minimal thyroid tilt/thick folds |
| | The | Glottal fry/creak |
| | Clouds have made a | Speech and cry with lowered larynx |
| | Way | Glottal fry/creak |
| | For you | Speech and cry with lowered larynx |
| | It's okay now | Speech and cry |

| Section | Lyrics | Voice quality |
|---|---|---|
| | You | Falsetto and cry |
| | Can laugh again | Speech and cry with larynx in a mid-position |

This table was used to aid the analysis of the results. Each participant's booklet was analysed in great detail and the points in the song where the squiggles represented peaks and troughs in emotional responses were noted. These peaks and troughs displayed by each individual were then mapped with regards to the particular voice qualities used at those moments in the song. For each individual a colour key was utilised to depict the intensity of the emotional responses experienced at different moments in the song by each individual. This is displayed in the section of the analysis in table 2. Refer to Appendix D for the complete table.

**Table 2- Colour key chart**

| | |
|---|---|
| | High intensity of emotional response |
| | Medium intensity of emotional response |
| | Low intensity of emotional response |

| Section | Lyrics | Voice quality | Sub - A | Sub - B | Sub - C | Sub - D | Sub - E | Sub - F | Sub - G | Sub - H |
|---|---|---|---|---|---|---|---|---|---|---|
| **Verse 1** | All a- | Breathy speech | | | | | | | | |
| | -lone inside four cold walls | Falsetto and cry | | | | | | | | |

| Section | Lyrics | Voice quality | Sub - A | Sub - B | Sub - C | Sub - D | Sub - E | Sub - F | Sub - G | Sub - H |
|---|---|---|---|---|---|---|---|---|---|---|
| | She's lost with just old memories | Breathy speech | | | | | | | | |
| | E- | Glottal fry/creak | | | | | | | | |
| | -motion- | Speech and cry | | | | | | | | |

This table highlights those voice qualities that evoked the strongest and weakest emotional reactions. In addition, the table presents patterns in the participant's responses each time specific qualities were featured. While an emphasis was placed on obtaining qualitative data in the methodology chapter, as the analysis of results progressed, the following method proved useful in gaining detailed insight into the intensity of emotional response evoked by each specific quality.

| Voice Quality | Number of opportunities to elicit a response | Number of high intensity emotional responses | Percentages |
|---|---|---|---|
| Breathy Speech | 40 | 13 | 32.5% |
| Falsetto and cry | 80 | 48 | 60% |
| Glottal Fry | 80 | 22 | 27.5% |
| Speech and cry | 128 | 48 | 37.5% |
| Speech and cry with a lowered larynx | 64 | 32 | 50% |
| Speech with minimal thyroid | 8 | 1 | 12.5% |

| Voice Quality | Number of opportunities to elicit a response | Number of high intensity emotional responses | Percentages |
|---|---|---|---|
| Speech and cry (thicker folds) | 64 | 26 | 40.625% |
| Belt | 16 | 6 | 37.5% |
| Speech and cry (thinner folds) | 8 | 2 | 25% |
| Raised vocal fold plane | 8 | 5 | 62.5% |

**Table 3- Table of opportunities**

### *4.1 Research Findings*

Table 3 indicates the number of high intensity emotional responses that were evoked by each quality. Each time a quality featured; there was the opportunity for eight possible responses in total. The amount of times each quality was heard has been multiplied by eight to give the total amount of opportunities for emotional response presented by each quality. The third column shows the number of times a high intensity emotional response was elicited by each quality. This information assisted the in-depth analysis connecting specific voice qualities to emotional responses outlined below.

Breathy speech featured several times throughout the composition and the emotional responses to it varied in intensity. It appeared that participants experienced a more intense emotional response to this quality as the composition developed. Four out of the eight participants experienced a

high intensity emotional response to breathy speech when it occurred towards the end of the first chorus. In addition, the same pattern of results was presented when this quality was heard in the second chorus. Overall, however, breathy speech evoked a fairly low emotional reaction from listeners, with only thirteen out of a possible forty opportunities presenting a high intensity emotional response.

There was a substantial emotional response displayed the first time that falsetto and cry quality appeared in the first verse and the intensity of responses increased as the composition developed, much as it did with breathy speech. Five out of the eight participants displayed squiggles that highlighted an intense emotional response to this quality towards the end of verse one. The reaction to falsetto and cry in the first chorus was almost unanimously high, with seven participants showing a clear increase in their emotional responses. Its appearance in the second chorus and middle eight section mimics a similar reaction. Falsetto and cry quality is heard again in the last chorus, however, this appeared to have less of an impact overall. Nevertheless, high intensity emotional responses were evoked by this quality over fifty per cent of the time that it was featured.

During the first bridge, speech and cry quality presented a low intensity emotional response, with only one participant producing a squiggle in an upward direction. However, this increased when the same quality was heard in the first chorus, with the majority of the participants appearing to

have some kind of emotional reaction. The intensity increased further at the end of the second chorus, resulting in seven participants displaying a significant emotional response to the quality. There was a drop off in intensity when speech and cry appeared in the middle eight, the last bridge and the final chorus. Overall, speech and cry had a fairly low emotional impact on the participants.

Speech and cry quality with the addition of a lowered larynx appeared frequently, though only lasted for short periods of time. This quality appeared to evoke fairly intense emotional reactions, with three of the participants displaying a significant increase in their emotional responses when this quality appeared in the first verse. This pattern of results was repeated in the first bridge, with again three participants displaying squiggles moving in an upward direction. Likewise, all three choruses and the middle eight section exhibited similar results. A high intensity emotional reaction was elicited by this quality in exactly half of times that it featured.

Speech quality with minimal thyroid tilt was heard in the first bridge, with a corresponding decline in intensity of emotional reaction amongst participants. Of the one opportunity for this quality to evoke an emotional reaction within the listener, only one participant displayed an increase in intensity of emotional response. These results strongly suggest that a lack of thyroid tilt could be the reason for this.

At the end of the first bridge, speech and cry quality appeared with thicker folds. The results displayed high intensity emotional responses, with two of the participants exhibiting a medium response and one participant displaying an intense emotional response. Furthermore, this quality appeared to evoke a very strong emotional reaction at the beginning of the second chorus, where all but one participant produced a squiggle moving in an upward direction. These results were consistent at the end of the second chorus, however, the intensity decreased in the third bridge and third chorus. Overall, this quality proved to evoke a fairly low emotional response with only twenty-six out of a possible sixty-four opportunities displaying an increase in intensity of emotional response.

Belt quality featured twice during the composition, firstly during the middle eight section. Here, belt appeared to have an impact on the intensity of listeners' responses, with all participants either displaying medium or high intensity emotional responses. The second occurrence of belt quality in chorus three had less of an impact, with a clear drop off in intensity amongst participants after hearing the quality.

When glottal fry first appeared in the first verse, three out of the eight participants exhibited a strong emotional response to the quality. When it was heard again, the response was not as strong, however, intensity of emotional responses increased later on in the first bridge for almost half of the participants. This quality received the strongest reaction when it

appeared in the second verse and during the third bridge, with half of the participants documenting an increase in intensity of emotional response. This quality had less of an impact during the last chorus. Overall this quality aroused fairly low intensity emotional reactions with only twenty-two out of eighty possible opportunities displaying an increase in emotional intensity experienced.

# 5.Evaluation

Falsetto quality with the addition of thyroid tilt elicited the highest intensity in emotional response overall. The addition of cry quality (thyroid tilt) presented a sweeter and more refined version of falsetto's purest form. High intensity emotional reactions to this quality may be attributed to the associations linked with it. For example, falsetto quality is often used when speaking or singing to children and its gentle nature gives it a sense of vulnerability. Additionally, it is possible that the listeners' responses to this quality were influenced by memories and past experiences through the mechanisms of evaluative conditioning and episodic memory (Juslin et al., 2010). On several occasions, this thesis has referred to the ability of cry quality to trigger in the listener the same laryngeal position that we access when we cry for real. This is an example of emotional contagion (Juslin, 2001) and could explain why this quality aroused the highest intensity emotional responses within the listeners. Therefore, it would have been interesting to note the emotional reactions experienced if this quality had appeared in its original form, with the absence of thyroid tilt. The presence of speech quality with a lack of thyroid tilt elicited the lowest emotional reactions overall. This further highlights the emotional impact that thyroid tilt (cry quality) has on the listener. Abrupt shifts of the vocal fold plane from horizontal to raised (speech to falsetto), giving a yodel effect, sparked high intensity emotional reactions in over half of the opportunities that were presented by this sound. This could be an example of musical

expectancy (Juslin et al., 2010) in action. It is likely that the sudden shift between the two qualities took the listener by surprise by presenting something that was different from expected.

Speech and cry quality with the addition of a lowered larynx evoked high intensity emotional reactions in exactly half of the opportunities that were presented for it to do so. This quality is a permutation of sob quality. During the composition, the singer repeatedly accessed a similar laryngeal posture to that of deep mourning or grief. Again this suggests the impact of emotional contagion (Juslin, 2001) on the listeners. Mimicking this unnaturally low position, even momentarily, could trigger the emotions mentioned above. In addition, the sound itself might have triggered a specific picture in the mind of the listener through the mechanism visual imagery (Juslin et al., 2010), resulting in an intense emotional reaction.

Overall, belt quality had less of an emotional impact than might have been expected. Excitement and surprise are commonly associated with this quality; however, not even half of the opportunities presented by belt quality elicited high intensity emotional responses from the listeners. This highlights the potential that taste has to influence emotional response. For some, this quality might not be considered aesthetically pleasing and this opinion could have indirectly influenced the participants' squiggles.

As the composition progressed, the vocalist used thicker vocal fold mass. The thicker the folds, the louder we perceive the sound to be. Although initial responses to speech and cry were fairly low intensity, as the mass of the folds increased during the development of the song, there was a corresponding increase in intensity of emotional reaction. This suggests that an increase in volume results in an increase in intensity of emotional response. Additionally, there is a certain amount of expectation attached to the development of a song, whereby it builds in intensity before reaching a climax. This cultural expectation could also have had an influence on the results displayed in this study.

Glottal fry (low level distortion) evoked high intensity emotional reactions in just over a quarter of the opportunities presented. This musical event is short and there is no pitch associated with it. As mentioned in the literature review, the incorporation of an unexpected distorted sound that opposes general aesthetic bias could take the listener by surprise and consequently result in an intense emotional reaction.

The listeners' responses intensified overall as the composition developed. Again this could be due to the innate expectation that as a song progresses, the intensity also increases. This might have influenced the reactions experienced. Similarly, this may be another example of emotional contagion (Juslin, 2001). Emotions experienced by the participants may have intensified as they imitated the changing emotional

state of the singer. Additionally, it is quite evident from the start that the general mood of the composition is solemn. Participants' unconscious detection of this may have set the foundations for the emotional responses that followed. Unfortunately, as mentioned previously, perceptive based studies rarely if ever draw solid conclusions. Part of what made the study problematic was the variables that could not be controlled. Human responses are complex and reactions to stimuli that bring about an emotional response are undoubtedly going to be affected by subconscious consultation with memories and what we individually believe to be true. This appears to be impossible to eliminate.

Although a variety of voice qualities and their permutations featured in this study, there are more sounds available to the human voice than were demonstrated in the composition. Additionally, some of the qualities that were defined in the methodology section were not used. These included speech, falsetto, cry and sob in their original forms. This is because, in the context of singing, different voice qualities are rarely used in a primary colour fashion. It is the mixing of these qualities that allows the singer to be artistic. Furthermore, the addition of thyroid tilt is what allows these qualities to sound like they are being sung rather than spoken, as well as enabling them to be taken higher in the range. Distortion in the form of false vocal fold constriction was not demonstrated simply because the stimulus did not cause the vocalist to respond in this way. Nevertheless, had all the qualities been incorporated, including those in their purest form,

this could have broadened the opportunity for emotional reaction. This highlights a potential flaw with the design of the stimulus.

### *5.1 Conclusion*

The chosen methodology worked well in that it allowed for spontaneous responses to be recorded in real time. However, the analysis process flagged up some issues surrounding the reliability of the data. For example, responses varied depending on where in the composition certain qualities were heard, this suggests that the context in which the voice qualities appear would ultimately have an impact on the listeners responses. Furthermore, the amount of times each quality featured in the composition was multiplied by the number of participants and noted alongside the number of intense emotional responses that it evoked. For example, breathy speech appeared five times in the composition, this was multiplied by eight participants to give forty potential responses. Out of those responses, thirteen were of a high intensity. This appeared to be a logical way of analysing the results. However, the amount of times each quality featured was not equal. That is to say that falsetto and cry quality appeared ten times within the piece, compared to belt quality that only appeared twice. This means that falsetto and cry essentially had more opportunity to elicit an intense emotional response. Each quality could have been featured an equal amount of times in the composition, however, this would have been quite an unnatural process for the vocalist. The voice qualities used were text driven as the vocalist responded to the stimulus, which was the lyrical content. Had she been told to use each

quality an equal amount of times, she would not have been able to respond organically and the song would not have communicated as well. Another issue that arose from the analysis process was the difficulty in making sense of each participant's squiggle. This was first highlighted in the second pilot study. Some participants produced clear peaks and troughs, whereas other illustrations displayed a generalised wave, which made it more difficult to distinguish at which points they were emotionally affected. As the objective was to allow participants to respond spontaneously and mindlessly to what they were hearing, they would probably have difficulty attempting to explain their squiggles themselves. As a result, analysis of the squiggles was heavily based on the author's interpretation and didn't necessarily reflect what the participant had intended. Additionally, the construction of the colour coded chart shown in table two, was once again heavily reliant on the ability of the author to interpret each participant's squiggle. Three colours were used: yellow for low intensity responses; orange for medium intensity responses and red for high intensity responses. It was down to the author to decide which category the responses belonged to. This makes the reliability of the examination of the data questionable. Without altering the design of the methodology, problems surrounding interpretation remain. Introducing other shades of colour along the yellow/orange/red spectrum may offer more options in terms of categorising each response more specifically. This could present a more detailed way of interpreting the results and improve reliability.

In the same way that individual perception could not be controlled, the emotional state of the participants prior to the study commencing presented another variable that had not been considered initially. For example, if one of the participants had received some upsetting news before walking into the studio, one would assume that their present mood could affect their emotional response to the stimulus. Questioning each individual on how they are feeling before the study starts could in some ways improve reliability of the results. Perhaps only individuals that were in a neutral emotional state should participate. This of course assumes that the participants would willingly offer such information and could also be perceived as quite invasive. Additionally, this brings us back to the problems posed by perception. Asking participants how they feel means relying on their ability to accurately articulate so.

With regards to the question stated at the beginning of this thesis, this study highlighted that different voice qualities most certainly have the ability to evoke an emotional response within the listener. What the study did not decipher was the specific emotions evoked as a result of particular qualities and moreover, whether these responses were a direct result of the voice quality alone. This is because the study did not eliminate variables such as lyrical content, melody and individual associations, all of which have the ability to massively influence emotional response. More research is required whereby these variables are controlled. This study could be repeated using the speaking voice rather than the singing voice.

Actors could deliver the text, responding to the stimulus through the use of different voice qualities, the only difference being that the melody is removed, eradicating one of the variables. Nevertheless, the main issue in terms of being able to explore this question more thoroughly lies with the inability to successfully control perception. Participants could be asked to respond to standardised examples of different voice qualities like those demonstrated in the accompanying audio files. Presenting the qualities as absolutes and providing prior training to participants, much like Colton and Estill did in their perceptual study (Colton and Estill, 1981), could lessen some of the problems surrounding perception. In order to implement this, the design of methodology would ultimately need to change to a scale of some description, whereby participants are able to rate intensity of emotional response to the individual qualities. This brings us full circle to the first approach considered in the methodology chapter. In addition, measuring physiological reactions to emotion, an option considered prior to designing this study could eliminate a certain amount of participant interference, though the caveat to this is that its invasive nature might actually hinder the reliability of the results.

In conclusion, it appears that the research idea was better than the methodology chosen. This is highlighted in the challenges that arose in gathering and analysing the data reliably. What is certain is that the participants were emotionally affected by the stimulus. Whether this was solely down to the use of different voice qualities cannot be ascertained as

other variables remained present. Focusing on measuring intensity of emotional response gave the study direction but also inadvertently changed the focus of the study whereby the research question could more appropriately have read, *which voice qualities evoke the most intense emotional responses within the listener?* I feel that I may have moved closer to giving an absolute answer to my question had more emphasis been placed on eradicating those remaining variables that could have an impact on emotional response. An adapted version of the research question is required to explore this topic more thoroughly, perhaps this could be, *Is it possible for emotional responses to musical stimuli to be evoked through the use of different voice qualities alone?* Nevertheless, the study itself served the current question and the trends displayed in the results supported some of the ideas and concepts researched and outlined in the literature review. Moreover, as a singer and singing teacher, this process has enabled me to learn more about the production of the different voice qualities at our disposal and therefore helped me to improve as both a performer and in particular, with regards to my coaching methods. A real bonus to the research is that I now know how to best articulate to my students what muscles and structures I want them to manipulate to produce a given sound. This makes the research extremely purposeful. I still maintain what I could describe as the initial driving force behind this research, that as performers we hope to affect and move our audience. At the beginning of this research I didn't know how this was possible; I just knew that it sometimes happened. As a result of exploring

the research question, I now have a better understanding of how and why the singer has the ability to emotionally move their audience.

### *5.2 Further research*

This study revealed that different voice qualities have the potential to elicit an emotional response within the listener. However, the context of which the quality is heard determines the responses of the listeners. Further research could include exploring the emotional impact that producing these different qualities may have in a therapeutic setting. Music therapy proves to be successful in enabling patients to increase their experiences of positive emotions and to reduce negative emotions through the use of music (Jones, 2005). The incorporation of producing different voice qualities in this setting may enhance positive results.

As a result of the extensive research conducted by Jo Estill, expert practitioners of her model have the information required to very specifically teach the production of different voice qualities. As part of a music therapy session, patients could be instructed on how to access qualities such as cry and belt, with the view of enabling them to release suppressed emotions that may be postponing their recovery. The act of tilting the thyroid cartilage forward for cry quality replicates the laryngeal posture for crying in everyday life. For those patients who are struggling to release feelings of grief and distress, this could be a powerful tool. This is due to the fact that what we do muscularly has the ability to affect our emotions

(Hamilton, 2015). As highlighted throughout this thesis, consciously putting the larynx into the cry posture could potentially trigger the urge to cry. Belt quality is the loudest sound that we can make. When we produce this quality we are accessing the yell posture. There is something quite therapeutic about yelling, possibly because of the physical commitment required and as a result of the intensity of the sound itself. Access to this quality may assist patients in releasing feelings of stress, frustration and anger.

In addition to aiding patient recovery, future research could potentially help to reveal some of the unanswered questions that arose from this study. Utilising different voice qualities in a therapeutic setting may help to discover which qualities trigger *specific* emotional responses. Essentially, giving the research a purposeful context could result in more specific results regarding the power of voice quality to evoke an emotional response not just within the listener, but also in the person producing the quality.

# 6. Appendices

## *6.1 Appendix A : Physiology and anatomy of the voice*

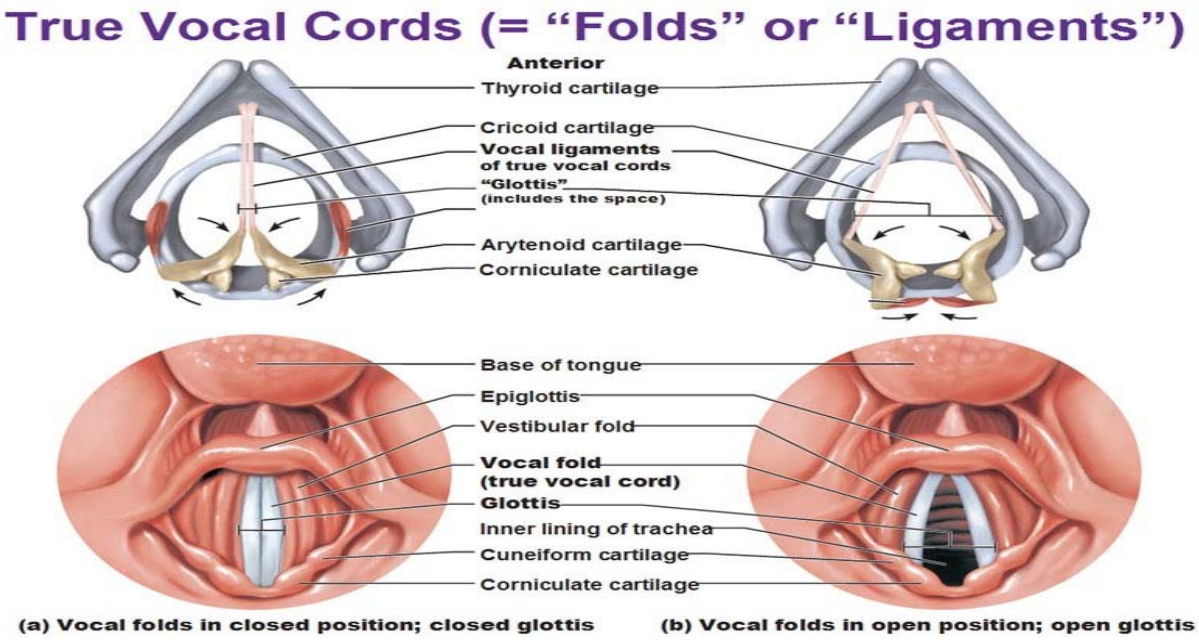
True vocal folds

The true vocal folds have both a body and a cover. The body consists of muscle and ligament and the cover is skin and pliant. It is the relationship between the body and cover that influences volume and the production of different voice qualities. (Obert and Chicurel, 2005). The true vocal folds originate at the thyroid cartilage and attach to the arytenoid cartilages (Obert and Chicurel, 2005). The muscles that act on these cartilages control the opening and closing of the true vocal folds (Obert and Chicurel, 2005). In order to produce voice, the true vocal folds must vibrate. This can happen when they are close enough for the passing breath to suck them together and blow them apart (Bernoulli's Principle) ((A) Klimek, et *al* 2005). Air pressure builds up beneath the vocal folds when they are closing or closed. The greater the pressure beneath the folds, the more intense the sound pressure wave and the louder the sound ((A) Klimek, et *al* 2005).

False vocal folds (Ventricular folds/Vestibular folds)

The false vocal folds are located above and to the side of the true vocal folds and approximate during protective closure (swallowing) and for thoracic fixation i.e. preparing to cough, sneeze or wherever strenuous activity occurs (Obert and Chicurel, 2005). Approximation of the ventricular folds is imperative to life; however, false vocal fold phonation is usually undesirable. This is because when the false vocal folds move in toward

one another, they impede on the free movement of the true vocal folds (Obert and Chicurel, 2005). This can be detrimental to vocal health; however, false vocal fold constriction can be used to produce a deliberately distorted sound.




**Glover, A, (no date)**

<u>Aryepiglottic Sphincter</u>

The Aryepiglottic Sphincter is a ring of muscle that originates at the arytenoid cartilages and inserts into the epiglottis. (Obert and Chicurel, 2005). When this muscle narrows it provides a formant between 2-4kHz. This bandwidth corresponds to the dimensions of the outer ear and is the area in which we hear best. As a result, the sound rings within the ear of the listener and it is perceived as piercing and loud ((A) Klimek, et *al* 2005). AES narrowing is the defining feature of twang quality ((A) Klimek,

et *al* 2005) and also features in other loud voice qualities such as belt and opera ((B) Klimek, et *al* 2005).

## *6.2 Appendix B : Instructions for the participants*

In front of you is a booklet with lyrics on, the purpose of this is to give you some kind of guideline when listening to the composition i am going to play you, try not to think about the words too much. Below the lyrics is a box, the bottom of the box will present a low intensity of emotion and the top of the box will show a high intensity of emotion.

When i play the composition i would like you to keep your pencil on the paper with in the box and allow your hand to express the intensity of the emotion you feel. If the line goes towards the upper end of the box the emotional response is more intense than the lower end of the box.

As result you will create a wave type format that will display your emotional response to the stimuli.

The composition consists of a solo voice, when listening i would like you to concentrate on the various voice qualities the singer uses and how intense you feel your emotional response is.

The purpose of this study is to establish wether or not different voice qualities evoke the intensity of the emotional response with in the listener.

Try not to over think it and let your hand move as unconsciously as possible.

ANY QUESTIONS?

### *6.3 Appendix C: Squiggle booklets*

Refer to accompanying folder for squiggle booklets

## *6.4 Appendix D: Description of voice qualities table*

### Table 1

| Section | Lyrics | Voice quality |
|---|---|---|
| **Verse 1** | All a- | Breathy speech |
| | -lone inside four cold walls | Falsetto and cry |
| | She's lost with just old memories | Breathy speech |
| | E- | Glottal fry/creak |
| | -motion- | Speech and cry |
| | -al she has so much to | Falsetto and cry |
| | say | Breathy speech |
| | Po- | Speech and cry with lowered larynx |
| | -wers bound her freedoms taken away | Breathy speech |
| **Bridge 1** | But it's okay now you can smile again | Speech quality with minimal thyroid tilt/thick folds |
| | The | Glottal fry/creak |
| | Clouds have made a | Speech and cry with lowered larynx |
| | Way | Glottal fry/creak |
| | For you | Speech and cry with lowered larynx |
| | It's okay now | Speech and cry |
| | You | Falsetto and cry |
| | Can laugh again | Speech and cry with larynx in a mid-position |
| | The | Glottal fry/creak |
| | Stars have made a place for you | Speech and cry with thicker folds |
| **Chorus 1** | Now the a- | Speech and cry |

| Section | Lyrics | Voice quality |
|---|---|---|
| | a-a- | Falsetto and cry (raised vocal fold plane) |
| | an-gels will | Breathy speech |
| | Hold you | Speech and cry with lowered larynx |
| | Now the a- | Speech and cry |
| | a-a- | Falsetto and cry |
| **Verse 2** | an-gels will hold you | Speech and cry with lowered larynx |
| | Star- | Glottal fry/creak |
| | -ing out wishing to be heard | Speech and cry with thicker folds |
| | Stor- | Glottal fry/creak |
| | -ies being told through her eyes | Speech and cry with thicker folds |
| | Ex- | Glottal fry/creak |
| | -pressionless she has so much to say | Speech and cry with thicker folds |
| | Powers bound her freedoms taken away | Speech and cry with thinner folds |
| **Bridge 2** | But its okay now you can smile again, the clouds have made a way for you | Speech and thyroid tilt but folds are still thick |
| | It's okay now you can laugh again, the stars have made a place for you | Speech and cry with thicker folds |
| **Chorus 2** | Now the angels will | Speech and cry with thicker folds |
| | Hold you | Speech and cry with lowered larynx |
| | Now the a-a- | Speech and cry |
| | -an- | Falsetto and cry |

| Section | Lyrics | Voice quality |
|---|---|---|
| | -gels will hold you | Speech and cry |
| **Middle eight** | Yeah | Falsetto that changes to speech and cry |
| | They will | Speech and cry |
| | ho- | Falsetto and cry |
| | -old you | Speech and cry with lowered larynx |
| | The angels | Falsetto and cry |
| | They will ho- | Belt |
| | -old you | Speech and cry with lowered larynx |
| **Bridge 3** | But its okay now you can smile again the clouds have made a way for you | Speech and cry |
| | It's okay now you can laugh again the | Speech and cry |
| | Stars | Glottal fry/creak |
| | have made a place for you | Speech and cry with thicker folds |
| **Chorus 3** | Now the a- | Speech and cry |
| | -a- an | Falsetto and cry |
| | -gels will hold you | Speech and cry |
| | Now the a- | Speech and cry |
| | -a-an | Falsetto and cry |
| | -gels will hold you | Speech and cry |
| | Now | Belt |

| Section | Lyrics | Voice quality |
|---|---|---|
| | They’ll hold you | Speech and cry |
| | Now | Glottal fry/creak |
| | The angels | Falsetto and cry |
| | Will | Glottal fry/creak |
| | Hold you | Speech and cry |

### *6.5 Appendix E: Colour key chart*

**Table 2**

Key

High intensity of emotional response

Medium intensity of emotional response

Low intensity of emotional response

**Table 2**

| Section | Lyrics | Voice quality | Sub - A | Sub-B | Sub-C | Sub-D | Sub-E | Sub-F | Sub-G | Sub-H |
|---|---|---|---|---|---|---|---|---|---|---|
| **Verse 1** | All a- | Breathy speech | | | | | | | | |
| | -lone inside four cold walls | Falsetto and cry | | | | | | | | |
| | She's lost with just old memories | Breathy speech | | | | | | | | |
| | E- | Glottal fry/creak | | | | | | | | |
| | -motion- | Speech and cry | | | | | | | | |
| | -al she has so much to | Falsetto and cry | | | | | | | | |
| | say | Breathy speech | | | | | | | | |

| Section | Lyrics | Voice quality | Sub - A | Sub-B | Sub-C | Sub-D | Sub-E | Sub-F | Sub-G | Sub-H |
|---|---|---|---|---|---|---|---|---|---|---|
| | Po- | Speech and cry with lowered larynx | | | | | | | | |
| | -wers bound her freedoms taken away | Breathy speech | | | | | | | | |
| **Bridge 1** | But it's okay now you can smile again | Speech quality with minimal thyroid tilt/thick folds | | | | | | | | |
| | The | Glottal fry/creak | | | | | | | | |
| | Clouds have made a | Speech and cry with lowered larynx | | | | | | | | |
| | Way | Glottal fry/creak | | | | | | | | |
| | For you | Speech and cry with lowered larynx | | | | | | | | |
| | It's okay now | Speech and cry | | | | | | | | |
| | You | Falsetto and cry | | | | | | | | |

| Section | Lyrics | Voice quality | Sub - A | Sub-B | Sub-C | Sub-D | Sub-E | Sub-F | Sub-G | Sub-H |
|---|---|---|---|---|---|---|---|---|---|---|
| | Can laugh again | Speech and cry with larynx in a mid-position | | | | | | | | |
| | The | Glottal fry/creak | | | | | | | | |
| | Stars have made a place for you | Speech and cry with thicker folds | | | | | | | | |
| **Chorus 1** | Now the a- | Speech and cry | | | | | | | | |
| | -an- | Vocal fold plane is abruptly raised into Falsetto | | | | | | | | |
| | -gels will | Breathy speech | | | | | | | | |
| | Hold you | Speech and cry with lowered larynx | | | | | | | | |
| | Now the a- | Speech and cry | | | | | | | | |
| | a- | falsetto and cry | | | | | | | | |
| | an-gels will hold you | Speech and cry with lowered larynx | | | | | | | | |

| Section | Lyrics | Voice quality | Sub - A | Sub-B | Sub-C | Sub-D | Sub-E | Sub-F | Sub-G | Sub-H |
|---|---|---|---|---|---|---|---|---|---|---|
| **Verse 2** | Star- | Glottal fry/creak | | | | | | | | |
| | -ing out wishing to be heard | Speech and cry with thicker folds | | | | | | | | |
| | Stor- | Glottal fry/creak | | | | | | | | |
| | -ies being told through her eyes | Speech and cry with thicker folds | | | | | | | | |
| | Ex- | Glottal fry/creak | | | | | | | | |
| | -pressionless she has so much to say | Speech and cry with thicker folds | | | | | | | | |
| | Powers bound her freedoms taken away | Speech and cry with thinner folds | | | | | | | | |
| **Bridge 2** | But its okay now you can smile again, the clouds have made a way for you | Speech with thyroid tilt but folds are still thick | | | | | | | | |

| Section | Lyrics | Voice quality | Sub - A | Sub-B | Sub-C | Sub-D | Sub-E | Sub-F | Sub-G | Sub-H |
|---|---|---|---|---|---|---|---|---|---|---|
| | It's okay now you can laugh again, the stars have made a place for you | Speech and cry with thicker folds | | | | | | | | |
| **Chorus 2** | Now the angels will | Speech and cry with thicker folds | | | | | | | | |
| | Hold you | Speech and cry with lowered larynx | | | | | | | | |
| | Now the a- | Speech and cry | | | | | | | | |
| | -a | Falsetto and cry | | | | | | | | |
| | an-gels will hold you | Speech and cry | | | | | | | | |
| **Middle eight** | Yeah | Falsetto and cry that moves seamlessly into speech and cry | | | | | | | | |
| | They will | Speech and cry | | | | | | | | |
| | ho- | Falsetto and cry | | | | | | | | |

| Section | Lyrics | Voice quality | Sub - A | Sub-B | Sub-C | Sub-D | Sub-E | Sub-F | Sub-G | Sub-H |
|---|---|---|---|---|---|---|---|---|---|---|
| | -old you | Speech and cry with lowered larynx | | | | | | | | |
| | The angels | Falsetto and cry | | | | | | | | |
| | They will ho- | Belt | | | | | | | | |
| | -old you | Speech and cry with lowered larynx | | | | | | | | |
| **Bridge 3** | But its okay now you can smile again the clouds have made a way for you | Speech and cry | | | | | | | | |
| | It's okay now you can laugh again the | Speech and cry | | | | | | | | |
| | Stars | Glottal fry/creak | | | | | | | | |
| | have made a place for you | Speech and cry with thicker folds | | | | | | | | |
| **Chorus 3** | Now the a- | Speech and cry | | | | | | | | |

| Section | Lyrics | Voice quality | Sub - A | Sub-B | Sub-C | Sub-D | Sub-E | Sub-F | Sub-G | Sub-H |
|---|---|---|---|---|---|---|---|---|---|---|
| | -a | Falsetto and cry | | | | | | | | |
| | an-gels will hold you | Speech and cry | | | | | | | | |
| | Now the a- | Speech and cry | | | | | | | | |
| | -a-an- | Falsetto and cry | | | | | | | | |
| | -gels will hold you | Speech and cry | | | | | | | | |
| | Now | Belt | | | | | | | | |
| | They'll hold you | Speech and cry | | | | | | | | |
| | Now | Glottal fry/creak | | | | | | | | |
| | The angels | Falsetto and cry | | | | | | | | |
| | Will | Glottal fry/creak | | | | | | | | |
| | Hold you | Speech and cry | | | | | | | | |

## 7. Appendices reference list

Glover, A. (no date) [electronic print] Available at: http://andrewglover.net/sing/vocal-folds/ (Accessed: 12th May 2015).

Klimek, M., Obert, K. and Steinhauer, K. (2005a) *The Estill voice training system level one: Compulsory figures for voice control.* Estill Voice Training Systems International, LLC.

Klimek, M., Obert, K. and Steinhauer, K. (2005b) *The Estill voice training system level two: Figure combinations for six voice qualities.* Estill Voice Training Systems International, LLC.

Obert, K, B., Chicurel, S, R. (2005) *Geography of the voice: Anatomy of an adams apple.* Estill Voice Training Systems International, LLC.

## 8. Thesis reference list

Baumgartner, H. (1992). ‘Rememberance of things past: Music, autobiographical memory, and emotion’, *Advances in consumer research,* 19, pp.613-620.

*Radio One Live Lounge* (2013) Radio one, 19 September.

Beechler, K., Estill, J., Fujimura, Osamu. and Sawada, M. (1996) *Vocal fold physiology: Controlling complexity and chaos: Temporal perturbation and voice qualities.* Singular Publishing Group Inc.

Bunt, L., Pavlicevic, M. (2001) ‘Perspectives from music therapy’, in: Juslin, P and Sloboda, J. (Eds). *Music and emotion.* Oxford: Oxford University Press pp.181-201.

Chescheir, W. M., (1985) ‘Some implications of Winnicotts concept’, *Clinical social work journal,* pp.229-232.

Chicurel, S. R., Obert, K, B. (2005) *Geography of the voice: Anatomy of an adam’s apple.* 2nd edn*.* Estill Voice Training Systems International, LLC.

Christensen, L, B., Johnson, R, B., Turner, L, A. (2014) *Research methods, design, and analysis.* 12th edn. New Jersey: Pearson education Inc.

Clarke, E, F. (2005) *Ways of listening: An ecological approach to the perception of musical meaning.* Oxford: Oxford University Press.

Colton, R, H., Estill, J, A. (1981) 'Elements of voice quality: Perceptual, acoustic, and physiologic aspects', *The Estill papers,* pp.312-401.

Crawford, J. R., Henry, J. D. (2004) 'The positive and negative affect schedule (PANAS): Constructive validity, measurement properties and normative data in a large non-clinical setting', *British journal of clinical psychology,* 43, pp. 245-265.

De Bodt, M, S. (1997) 'Test-retest study of the GRBAS scale: Influence of experience and professional background on perceptual rating of voice quality', *Journal of voice,* 11(1), pp.74-80.

EnfieldCouncil (2014) *Squiggle game.* Available at: https://www.youtube.com/watch?v=4kAtRtF_zeE (Accessed: 20th May 2015).

Estill, J. (1988) 'Belting and classic voice quality: Some physiological differences', *Medical problems of performing artists,* 3, pp.37-42.

Ficken, T. (1976) 'The use of songwriting in a psychiatric setting', *Journal of music therapy,* 13(4), pp.164-172.

Gregory, A. (1996) ‘Cross cultural comparisons in the affective response to music’, *Psychology of music: the journal of the society for research in psychology of music and music education,* 24(1), pp.47-52.

Gwet, K, L. (2014) *The Handbook of Inter-rater reliability: The definitive guide to measuring the extent of agreement among raters .* 4th edn. Advanced analytics LLC: Gaithersburg USA.

Hamilton, D. (2015) *I love me: The science of self-love.* London: Hay House UK Ltd.

Hevner, K. (1936) ‘Experimental studies of the elements of expression in music’, *The American journal of psychology,* 48(2), pp.246-268.

Jones, J. D. (2005) ‘A comparison of songwriting and lyric analysis techniques to evoke emotional change in a single session with people who are chemically dependent’, *Journal of music therapy,* 42(2), pp.94-110.

Juslin, P. N. (2001) ‘Communicating emotion in music performance: A review and a theoretical framework’, in P. N. Juslin & J.A. Sloboda. (eds). *Music and emotion: Theory and research.* Oxford: Oxford University Press pp. 309-337.

Juslin, P. N., Liljestrom, S., Vastfjall, D. and Lundqvist, L. (2010) ‘How does music evoke emotion? Exploring the underlying mechanisms’, in P. N. Juslin & J. A. Sloboda. (eds). *Music and emotion: Theory and research.* Oxford: Oxford University Press pp. 605-637.

Juslin, P. N., Sloboda, J. A. (2001) ‘Psychological perspectives on music and emotion’, in P. N. Juslin & J. A. Sloboda. (eds). *Music and emotion: Theory and research.* Oxford: Oxford University Press pp. 71- 98.

Juslin, P. N., Västfjäll, D. (2008) ‘Emotional responses to music: The need to consider underlying mechanisms’, *Behavioral and brain sciences,* 31(5), pp.559- 575.

Kayes, G. (2004) *Singing and the Actor.* 2nd edn. London: A & C Black Publishers Limited.

Klimek, M., Obert, K. and Steinhauer, K. (2005a) *The Estill voice training system level one: Compulsory figures for voice control.* Estill Voice Training Systems International, LLC.

Klimek, M., Obert, K. and Steinhauer, K. (2005b) *The Estill voice training system level two: Figure combinations for six voice qualities.* Estill Voice Training Systems International, LLC.

Krieman, J. (1993) ‘Perceptual evaluation of voice quality: Review, tutorial, and a framework for future research’, *Journal of speech and hearing research,* 36(1), pp.21-35.

Menasha, W. (1991) ‘Vocal quality factors: Analysis, synthesis, and perception’, *The journal of the acoustical society of America,* 90(5), p.2394-2410.

Speed, A. (2010) ‘Speech and cry’. *The advanced course in the Estill model, applications, combinations and variations*. Unpublished. The Voice Explained.

Tanner, B. A. (2011) ‘Validity of global physical and emotional SUDS’, *Applied psychophysiology and biofeedback,* 37(1), pp.31-34.

Taylor Swift, 2012. *Red.* [CD] USA: Big machine records.

The Mcgraw-Hill Companies, Inc. (2013) [Electronic print] Available at: http://www.lookfordiagnosis.com/mesh_info.php?term=Laryngeal+Cartilages&lang=1 (Accessed: 20 January 2015).

Tyson, E.H. (2002) ‘Hip hop therapy an exploratory study of rap music intervention with at-risk and delinquent youth’, *Journal of poetry therapy,* 15(3), pp.131-144.

Winnicott, D, W. (1958) *Through Pediatrics to psychoanalysis.* London: Hogarth Press.